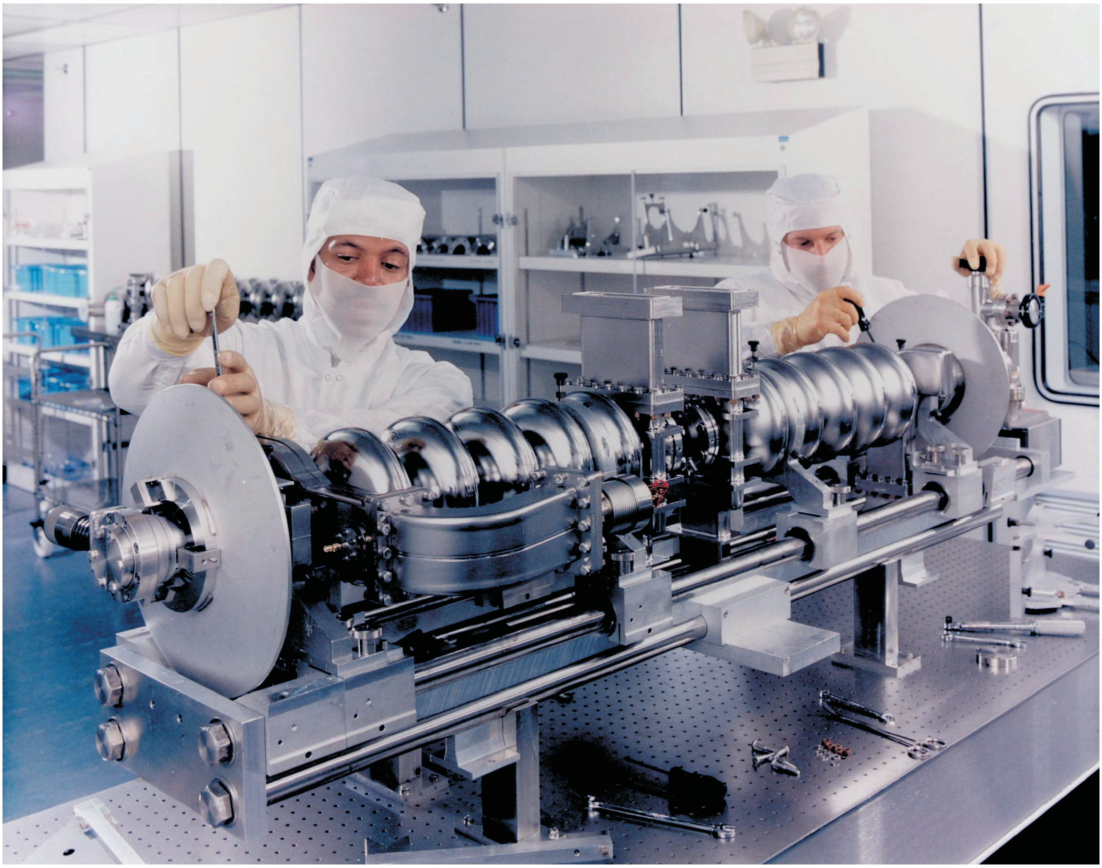
IZDAJA DRUŠTVO MATEMATIKOV, FIZIKOV IN ASTRONOMOV SLOVENIJE

ISSN 0473-7466

2008
Letnik 55
2

OBZORNIK ZA MATEMATIKO IN FIZIKO



OBZORNIK MAT. FIZ. • LJUBLJANA • LETNIK 55 • ŠT. 2 • STR. 41–80 • MAREC 2008

OBZORNIK ZA MATEMATIKO IN FIZIKO

Glasilo Društva matematikov, fizikov in astronomov Slovenije
Ljubljana, MAREC 2008, letnik 55, številka 2, strani 41–80

Naslov uredništva: DMFA–založništvo, Jadranska ulica 19, p. p. 2964, 1001 Ljubljana **Telefon:** (01) 4766 553, 4232 460 **Telefaks:** (01) 4232 460, 2517 281 **Elektronska pošta:** Zaloznistvo@dmfa.si **Internet:** http://www.obzornik.si/ **Transakcijski račun:** 03100–1000018787 **Devizna nakazila:** SKB banka d.d., Ajdovščina 4, 1513 Ljubljana **SWIFT (BIC):** SKBASI2X **IBAN:** SI56 0310 0100 0018 787

Jezikovno pregledal Janez Juvan.
Natisnila tiskarna COLLEGIUM GRAPHICUM v nakladi 1300 izvodov.

Člani društva prejemajo Obzornik brezplačno. Celoletna članarina znaša 21,00 EUR, za druge družinske člane in študente pa 10,50 EUR. Naročnina za ustanove je 33,38 EUR, za tujino 30 EUR. Posamezna številka za člane stane 4,18 EUR, stare številke 2,17 EUR.

DMFA je včlanjeno v Evropsko matematično društvo (EMS), v Mednarodno matematično unijo (IMU), v Evropsko fizikalno društvo (EPS) in v Mednarodno združenje za čisto in uporabno fiziko (IUPAP). DMFA ima pogodbo o recipročnosti z Ameriškim matematičnim društvom (AMS).

Revija izhaja praviloma vsak drugi mesec. Sofinancirata jo Javna agencija za raziskovalno dejavnost Republike Slovenije ter Ministrstvo za šolstvo in šport.

Poštnina plačana pri pošti 1102 Ljubljana

NAVODILA SODELAVCEM OBZORNIKA ZA ODDAJO PRISPEVKOV

Revija Obzornik za matematiko in fiziko objavlja izvirne znanstvene in strokovne članke iz matematike, fizike in astronomije, včasih tudi kak prevod. Poleg člankov objavlja prikaze novih knjig s teh področij, poročila o dejavnosti Društva matematikov, fizikov in astronomov Slovenije ter vesti o drugih pomembnih dogodkih v okviru omenjenih znanstvenih ved. Prispevki naj bodo zanimivi in razumljivi širšemu krogu bralcev, diplomantov iz omenjenih strok.

Članek naj vsebuje naslov, ime avtorja (oz. avtorjev), sedež institucije, kjer avtor(ji) dela(jo), izvleček v slovenskem jeziku, naslov in izvleček v angleškem jeziku, klasifikacijo (MSC oziroma PACS) in citirano literaturo. Slike in tabele, ki naj bodo oštevilčene, morajo imeti dovolj izčrpen opis, da jih lahko večinoma razumemo tudi ločeno od besedila. Avtorji člankov, ki želijo objaviti slike iz drugih virov, si morajo za to sami priskrbeti dovoljenje (copyright). Prispevki so lahko oddani v računalniški datoteki PDF ali pa natisnjeni enostransko na belem papirju formata A4. Zaželena velikost črk je 12 pt, razmik med vrsticami pa vsaj 18 pt.

Prispevke pošljite odgovornemu uredniku ali uredniku za matematiko oziroma fiziko na zgoraj napisani naslov uredništva. Vsak članek se praviloma pošlje dvema anonimnima recenzentoma, ki morata predvsem natančno oceniti, kako je obravnavana tema predstavljena, manj pomembna pa je originalnost (in pri matematičnih člankih splošnost) rezultatov. Če je prispevek sprejet v objavo, potem urednik prosi avtorja še za izvorne računalniške datoteke. Le-te naj bodo praviloma napisane v eni od standardnih različic urejevalnikov TEX oziroma LATEX, kar bo olajšalo uredniški postopek.

Avtor se z oddajo članka strinja tudi z njegovo kasnejšo objavo v elektronski obliki na internetu.

INTEGRALI ELEMENTARNIH FUNKCIJ

MARKO SLAPAR

Inštitut za matematiko, fiziko in mehaniko
Ljubljana

Math. Subj. Class. (2000): 12H05

Dobro znano dejstvo je, da za nekatere elementarne funkcije ne obstajajo nedoločeni integrali, ki bi se lahko zopet izražali samo s pomočjo elementarnih funkcij. Primer takega integrala je $\int e^{x^2} \mathrm{d}x$. V članku predstavimo, kaj je v ozadju te teorije, in pokažemo nekaj primerov elementarnih funkcij, ki nimajo elementarnih integralov.

INTEGRALS OF ELEMENTARY FUNCTIONS

It is a well known fact that certain integrals of elementary functions cannot be expressed in elementary terms. An example of such an integral is $\int e^{x^2} \mathrm{d}x$. In this paper we show what is behind this theory and give some examples of elementary functions that do not have elementary integrals.

1. Uvod

Iz osnovnega izreka integralskega računa vemo, da sta si računanje nedoločenega integrala in računanje odvoda bolj ali manj inverzni operaciji. Čeprav je računanje odvoda elementarne funkcije preprosto, pa to ne velja za računanje nedoločenega integrala. Med drugim je problem v tem, da se nedoločeni integral bolj nerodno obnaša na produktu funkcij. Metoda per partes nam, razen v izjemnih primerih, ne da prav veliko. Zato verjetno vsak, ki se je kdaj mučil z integracijo, sluti, da so elementarne funkcije, ki imajo elementarne nedoločene integrale, morda bolj izjema kot pravilo. V nadaljevanju članka bomo videli, da je ta slutnja seveda upravičena, vendar pa dokaz tega ni tako preprost.

Za začetek si poglejmo preprost primer integrala racionalne funkcije

$$\begin{aligned}\int \frac{x^4 - x^3 + x^2 + x - 1}{x^2(x^2+1)} \,\mathrm{d}x &= \int 1 \,\mathrm{d}x - \int \frac{1}{x^2} \,\mathrm{d}x + \int \frac{\mathrm{d}x}{x} + \int \frac{-2x+1}{x^2+1} \,\mathrm{d}x = \\ &= x + \frac{1}{x} + \log x - \int \frac{\mathrm{d}(x^2+1)}{x^2+1} + \int \frac{\mathrm{d}x}{x^2+1} = \\ &= \frac{x^2+1}{x} + \log \frac{x}{x^2+1} + \operatorname{arctg} x + C\,.\end{aligned}$$

Pri računanju integrala smo si pomagali z razcepom racionalne funkcije na parcialne ulomke [6, str. 237–239]. Ta metoda integracije racionalnih funkcij se včasih imenuje tudi metoda nedoločenih koeficientov in nam omogoča integracijo poljubne racionalne funkcije pod pogojem, da znamo imenovalec faktorizirati na nerazcepne faktorje. Ti so bodisi linearni bodisi kvadratni nerazcepni faktorji. V rezultatu integracije bodo v splošnem nastopale racionalne funkcije, logaritmi racionalnih funkcij in funkcije arctg, ki so posledica nerazcepnih kvadratičnih faktorjev. Slednjih se lahko znebimo, če smo pripravljeni v metodo vpeljati kompleksna števila, saj iz razcepa $x^2+1=(x+i)(x-i)$ sledi

$$\int \frac{\mathrm{d}x}{x^2+1} = \frac{i}{2}\int \frac{\mathrm{d}x}{x+i} - \frac{i}{2}\int \frac{\mathrm{d}x}{x-i}\,.$$

Integrala na desni strani sta seveda integrala kompleksnih funkcij. S pomočjo kompleksnega logaritma ($\log z = \log|z| + i\mathrm{Arg}\ z$), ima sedaj rezultat obliko

$$\int \frac{x^4-x^3+x^2+x-1}{x^3(x^2+1)}\,\mathrm{d}x = \frac{x^2+1}{x} + \log\frac{x}{x^2+1} + \frac{i}{2}\log\frac{x+i}{x-i} + C\,.$$

V splošnem lahko vsak polinom v kompleksnem razcepimo na same linearne faktorje. Vsako racionalno funkcijo tako lahko razcepimo na parcialne ulomke, ki bodo v imenovalcih imeli samo potence linearnih faktorjev. Tem sumandom lahko poiščemo nedoločene integrale, ki bodo bodisi zopet cele potence linearnih faktorjev ali pa kompleksni logaritmi linearnih faktorjev. Zato imamo izrek

Izrek 1 (Laplace, 1812). *Nedoločeni integral racionalne funkcije je vedno elementarna funkcija. Le-ta je ali racionalna funkcija ali pa vsota racionalne funkcije in logaritmov racionalnih funkcij, pomnoženih s konstantami.*

Nabor funkcij, ki jih dobimo v integralih racionalnih funkcij, se je z vpeljavo kompleksnih števil zmanjšal, poenostavila pa se je tudi sama metoda integracije racionalnih funkcij. Seveda pa je slaba stran, da bodo povsem realne funkcije dobile nedoločene integrale, ki bodo (navidezno) kompleksni, zato smo za zdaj seveda lahko skeptični glede smiselnosti vpeljave kompleksnih števil. Bolj prepričljive razloge bomo videli v nadaljevanju. Izkazalo se bo namreč, da se integrali elementarnih funkcij, če so seveda elementarni, v splošnem izražajo kot vsota funkcij „iste vrste“ in logaritmov le-teh.

Čeprav se kar nekaj tipov nedoločenih integralov lahko s primerno substitucijo prevede na integral racionalne funkcije, pa v splošnem zgodba ni tako preprosta. Medtem ko je odvod elementarne funkcije zopet elementarna funkcija, pa za integral to vsekakor ni res. Tako se na primer integrali

- $\int e^{-x^2}\mathrm{d}x$
- $\int \frac{\mathrm{d}x}{\log x}$
- $\int \frac{\sin x}{x}\,\mathrm{d}x$
- $\int \frac{\mathrm{d}x}{\sqrt{x^3+1}}$

ne morejo zapisati samo z elementarnimi funkcijami. Naj omenimo, da je prvi integral tesno povezan z distribucijo normalnih spremenljivk v verjetnosti, drugi z distribucijo praštevil, zadnji pa spada med eliptične integrale.

Prav tako je včasih težko oceniti, kdaj je integral funkcije elementaren in kdaj ne. Integrala $\int x^x \, \mathrm{d}x$ in $\int x^x \log x \, \mathrm{d}x$ nista elementarna, medtem ko je $\int (x^x + x^x \log x) \, \mathrm{d}x = x^x + C$ elementaren. Podobno je integral $\int \left(\frac{e^x}{x^n} + \frac{ae^x}{x}\right) \mathrm{d}x$ elementaren pri konstanti $a = -\frac{1}{(n-1)!}$ (per partes), drugače pa je neelementaren.

2. Elementarne funkcije

V tem razdelku bomo natančno opisali, kaj si predstavljamo pod pojmom elementarne funkcije. Pod pojmom elementarne funkcije si običajno predstavljamo kakršnekoli funkcije, ki jih dobimo iz funkcije $f(x) = x$ le s pomočjo znanih operacij in funkcij: seštevanja, odštevanja, deljenja, množenja, korenjenja, potenciranja, trigonometričnih funkcij in njihovih inverzov, eksponentne in logaritemske funkcije, ter poljubnih kompozitumov le-teh. Na primer

$$\frac{ex^3 - 7x \log x}{\sqrt[6]{e^{-x} - \sin(x/(x^5 + 6))}} \, .$$

Nabor funkcij, ki sestavljajo elementarne funkcije, se znatno zmanjša, če pri računanju dovolimo še kompleksna števila. V kompleksnem namreč velja

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i} \quad \text{in} \quad \cos x = \frac{e^{ix} + e^{-ix}}{2} \, . \tag{1}$$

S pomočjo teh formul lahko dobimo na primer

$$\arcsin x = -i \log\left(ix + \sqrt{1 - x^2}\right) \quad \text{in} \quad \operatorname{arctg} x = \frac{1}{2i} \log \frac{1 + ix}{1 - ix} \, .$$

Vse trigonometrične in inverzne trigonometrične funkcije se zato lahko zapišejo le s pomočo logaritemske in eksponentne funkcije, kar računanje in kasnejše definicije precej olajša.

Spomnimo se, da je funkcija f *analitična* na odprtem intervalu $I \subset \mathbb{R}$, če jo v vsaki točki iz I lahko razvijemo v Taylorjevo vrsto, ki bo na neki okolici te točke tudi konvergirala proti funkciji. Kompleksna funkcija $f = u + iv$ je analitična, če sta analitična tako njen realni del u kot njen imaginarni del v. Analitičnost se seveda ohranja pri seštevanju, odštevanju, množenju funkcij in tudi pri odvajanju ter integriranju. Kvocient analitičnih funkcij bo analitičen na komplementu ničel funkcije v imenovalcu. Naj bo f kompleksna analitična funkcija na odprtem intervalu I in naj tam nima ničel.

Točka x_0 naj bo poljubna točka na intervalu I. *Logaritem* funkcije f definiramo s predpisom $(\log f)(x) = \int_{x_0}^{x} f'(t)/f(t)\,\mathrm{d}t$, kjer nam izraz pomeni običajen Riemannov integral. Sprememba referenčne točke x_0 spremeni logaritem funkcije do aditivne konstante, kar pa nas ne bo motilo. Če želimo, lahko konstanto vedno določimo tako, da bo veljalo $e^{\log f} = f$. Funkcija $\log f$ je analitična funkcija, saj je integral analitične funkcije. Analogno bi lahko dejali, da je $\log f$ katerakoli (analitična) funkcija g, za katero velja $f'/f = g'$. Če pa je kdo bolj doma v kompleksni analizi, lahko za $\log f$ vzame tudi katerokoli vejo kompleksne logaritemske funkcije. Za poljubno analitično funkcijo f na intervalu I bomo rekli funkciji e^f *eksponentna funkcija* funkcije f. Le-ta je na I zopet analitična.

Za kompleksno funkcijo h rečemo, da je *meromorfna* na intervalu $I \subset \mathbb{R}$, če je kvocient dveh funkcij f in g ($g \not\equiv 0$), ki sta obe analitični na I. Prostor meromorfnih funkcij na $I \subset \mathbb{R}$ označimo z $\mathcal{M}(I)$ in je očitno obseg. Obseg $\mathcal{M}(I)$ bo za nas pomenil osnovni razred funkcij, s katerimi bomo računali.

Naj bo F poljuben obseg. Z $F[x]$ označimo kolobar polinomov nad obsegom F v nedoločenki x in z $F(x)$ obseg racionalnih funkcij nad F v eni spremenljivki. Analogno imamo v več spremenljivkah kolobar $F[x_1, x_2, \ldots, x_n]$ polinomov v n spremenljivkah in $F(x_1, x_2, \ldots, x_n)$ obseg racionalnih funkcij v n spremenljivkah. Če je $G \supset F$ razširitev obsega F in $\alpha \in G$, je α *algebraičen* element nad F, če je $p(\alpha) = 0$ za kakšen polinom $p \in F[x]$. V nasprotnem primeru je α *transcendenten* element nad F. Za $\alpha \in G$, transcendenten nad F, nam preslikava $\alpha \mapsto x$ poda izomorfizem tako kolobarjev $F[\alpha]$ in $F[x]$ kot tudi obsegov $F(\alpha)$ in $F(x)$. Več in bolj natančno lahko o razširitvah obsegov najdemo v [5].

Definicija 1. Naj bodo $f_1, f_2, \ldots, f_n$ meromorfne funkcije na intervalu I. Potem je $\mathbb{C}(f_1, f_2, \ldots, f_n)$ obseg funkcij oblike

$$h = \frac{p(f_1, f_2, \ldots, f_n)}{q(f_1, f_2, \ldots, f_n)}\,,$$

kjer sta $p, q \in \mathbb{C}[x_1, x_2, \ldots, x_n]$ polinoma in $q \neq 0$.

Definicija 2. *Elementaren obseg* funkcij je vsak obseg oblike $\mathbb{C}(x, f_1, \ldots, f_n)$, kjer so $f_1, \ldots, f_n$ meromorfne funkcije na nekem intervalu $I \subset \mathbb{R}$, in je f_k ali algebraična nad $\mathbb{C}(x, f_1, \ldots, f_{k-1})$, ali je logaritem funkcije iz $\mathbb{C}(x, f_1, \ldots, f_{k-1})$, ali pa je eksponentna funkcija neke funkcije iz $\mathbb{C}(x, f_1, \ldots, f_{k-1})$. Meromorfna funkcija f je *elementarna*, če je vsebovana v kakšnem elementarnem obsegu funkcij.

Čeprav je interval I, na katerem so funkcije meromorfne, formalno del definicije, pa se s tem ne bomo preveč obremenjevali. Princip enoličnosti nam namreč pove, da je meromorfna funkcija natanko določena že, ko poznamo njene vrednosti na katerikoli množici s stekališčem. Za interval lahko tako vzamemo katerikoli interval, na katerem so vse funkcije dobro definirane. V nadaljevanju bomo na interval I zato preprosto pozabili.

Primer 1. Funkcija $f(x) = x^s,\ s \in \mathbb{R}$, je elementarna funkcija, saj je $f \in \mathbb{C}(x, \log x, e^{s \log x})$. Če je $s \in \mathbb{Q}$, je f tudi algebraična nad $\mathbb{C}(x)$, in je zato f že v elementarnem obsegu $\mathbb{C}(x, f)$, če pa je s celo element $\mathbb{Z}$, je f že v $\mathbb{C}(x)$.

Primer 2. Iz formul (1) sledi, da so trigonometrične funkcije $\sin x$, $\cos x$ in $\tan x$ v elementarnem obsegu $\mathbb{C}(x, e^{ix})$, funkcija $\arcsin x$ pa se nahaja v elementarnem obsegu $\mathbb{C}(x, \sqrt{1-x^2}, \log(ix + \sqrt{1-x^2}))$.

Primer 3. Funkcija

$$\frac{\pi x^3 - 7x \ln x}{\sqrt[6]{e^{-x} - \sin(x/(x^5+6))}}$$

je v elementarnem obsegu $\mathbb{C}\Big(x, \ln x, e^x, e^{i\frac{x}{x^5+6}}, \sqrt[6]{e^{-x} - \sin(x/(x^5+6))}\Big)$.

Pomembna lastnost obsegov elementarnih funkcij je, da so zaprti za odvajanje.

Izrek 2. *Naj bo funkcija f v elementarnem obsegu K. Potem je $f' \in K$.*

Dokaz. Naj bo $K = \mathbb{C}(x, f_1, \ldots, f_n)$. Izrek bomo dokazali z indukcijo po n. Če je $n = 0$, je $K = \mathbb{C}(x)$. Ker je odvod racionalne funkcije racionalna funkcija, je K zaprt za odvajanje. Predpostavimo torej, da je $L = \mathbb{C}(x, f_1, \ldots, f_{n-1})$ zaprt za odvajanje in je $K = L(f_n)$, pri čemer je f_n ali logaritem, ali eksponent funkcije iz L, ali pa algebraičen nad L. Dovolj je, če pokažemo, da je $f_n' \in K$. Če je $f_n = e^g$, $g \in L$, potem je $f_n' = g'e^g$, in ker je $g' \in L$ po predpostavki, je $f_n' \in K$. Podobno naj bo $f_n = \log g$ in $f_n' = g'/g \in L \subset K$. Naj bo sedaj f_n algebraičen nad L. Predpostavimo, da je $p(f_n) = 0$ za polinom $p(t) = t^m + a_{m-1}t^{m-1} + \cdots + a_0$, kjer je stopnja polinoma p najmanjša možna. Zato velja $q(f_n) \neq 0$ za polinom $q(t) = mt^{m-1} + a_{m-1}(m-1)t^{m-2} + \cdots + a_1$. Po drugi strani pa velja

$$0 = (p(f_n))' = q(f_n)f_n' + a_{m-1}' f_n^{m-1} + a_{m-2}' f_n^{m-2} + \cdots + a_0'.$$

Torej je

$$f_n' = -\frac{a_{m-1}' f_n^{m-1} + a_{m-2}' f_n^{m-2} + \cdots + a_0'}{q(f_n)}.$$

Ker so odvodi $a_k' \in L$ po indukcijski predpostavki, je desna stran v K. S tem je dokaz končan. ∎

3. Liouvillov izrek

V prejšnjem razdelku smo pokazali, da je odvod vsake elementarne funkcije prav tako elementarna funkcija. Še več, izrek, ki smo ga pokazali zgoraj, nam pove, da odvod ni nič bolj kompliciran kot sama funkcija, v smislu, da se nahaja v istem elementarnem obsegu. Kot bomo videli, za integral ne velja enako. Integral elementarne funkcije je pogosto neelementaren. Če pa je integral elementarne funkcije elementaren, lahko precej natančno povemo, kakšno obliko bo imel. To je vsebina Liouvillovega izreka iz leta 1835, ki je osnova tako za algoritmično računanje nedoločenih integralov [2] kot tudi za dokazovanje, da določena elementarna funkcija nima elementarnega integrala. Izreka tukaj ne bomo pokazali, za dokaz pa se bralec lahko obrne na [3].

Izrek 3 (Liouville, 1835). *Naj bo $f \in K = \mathbb{C}(x, f_1, \ldots, f_n)$. Potem ima f elementaren integral natanko takrat, ko velja*

$$f = \sum_{k=1}^{m} a_k \frac{g_k'}{g_k} + h',$$

kjer so $a_1, \ldots, a_m \in \mathbb{C}$ in $g_1 \ldots, g_k, h \in K$. Takrat velja

$$\int f(x)\,\mathrm{d}x = \sum_{k=1}^{m} a_k \log g_k + h\,.$$

Naj bo K kot v formulaciji izreka in označimo s

$$K_0 = \{f \in K; f = \sum_{k=1}^{m} a_k \frac{g_k'}{g_k} + h',\ \text{kjer so } g_1, \ldots, g_m, h \in K,\ a_1, \ldots, a_m \in \mathbb{C}\}.$$

Liouvillov izrek nam pove, da ima elementarna funkcija f, ki se nahaja v elementarnem obsegu K, elementaren integral natanko tedaj, ko se nahaja v podmnožici K_0. Ker za dva elementarna obsega $L \subset K$ seveda velja $L_0 \subset K_0$, je vseeno, kateri elementarni obseg izberemo za preverjanje elementarnosti integrala funkcije f. Vsekakor pa je bolje vzeti čim „manjši“ obseg, saj bo tako preverjanje lažje. V primeru, ko je $K = \mathbb{C}(x)$ obseg racionalnih funkcij, nam Laplaceov izrek pove, da je $K_0 = K$, in ima tako vsaka racionalna funkcija elementaren integral. Prav tako je $K = K_0$ za elementarne obsege tipa $\mathbb{C}(x, \sqrt[n]{(ax+b)/(cx+d)})$ ali $\mathbb{C}(x, \sqrt{ax^2+bx+c})$, saj znamo integrale funkcij iz teh obsegov s primernimi substitucijami prevesti na integrale racionalnih funkcij. V splošnem pa bomo videli, da $K \backslash K_0 \neq \emptyset$, in s tem pokazali obstoj elementarnih funkcij, ki nimajo elementarnih integralov. Prva taka preprosta elementarna obsega sta $\mathbb{C}(x, e^g)$ in $\mathbb{C}(x, \log g)$, kjer je g poljubna nekonstantna racionalna funkcija.

Opomba 1. Ko integriramo povsem realno funkcijo, običajno želimo tudi končni rezultat predstaviti samo z realnimi funkcijami. To seveda lahko vedno storimo s pomočjo primernih formul, ki nam bodo kompleksne funkcije zamenjale z realnimi. Vendar pa je pri Liouvillovem izreku ključno, da razširimo razred funkcij na kompleksne. Primer je $\int(1/(1+x^2))\,\mathrm{d}x = \operatorname{arctg} x$, ki ga ne moremo zapisati kot vsoto realne racionalne funkcije in z realnimi konstantami pomnoženih logaritmov realnih racionalnih funkcij. To dejstvo sicer ni povsem očitno in je lepa vaja.

4. Elementarna obsega $\mathbb{C}(x, e^g)$ in $\mathbb{C}(x, \log g)$

Lema 4. *Če je g racionalna funkcija, ki ni konstanta, je $e^{g(x)}$ transcendentna nad $\mathbb{C}(x)$.*

Dokaz. Predpostavimo, da je e^g algebraična nad $\mathbb{C}(x)$. Potem obstaja monični polinom $p \in \mathbb{C}(x)[y]$ minimalne stopnje n, da velja

$$p(e^g) = e^{ng} + a_{n-1}e^{(n-1)g} + \cdots + a_0 = 0\,.$$

Odvajamo zgornjo enakost in dobimo novo polinomsko enačbo

$$q(e^g) = ng'e^{ng} + (a'_{n-1} + (n-1)g'a_{n-1})e^{(n-1)g} + \cdots + a'_0 = 0\,.$$

Ker je bil polinom p minimalen, mora p deliti polinom q. Posebej mora veljati $a'_0 = ng'a_0$ oziroma

$$\frac{a'_0}{a_0} = ng'\,.$$

Če enačbo integriramo, dobimo $\log a_0 = ng + C$ oziroma $a_0 = De^{ng}$, kjer sta C in D kompleksni konstanti. Funkcija e^g, razumljena kot funkcija, definirana na kompleksnih številih, ima v polih racionalne funkcije g bistveno singularnost. Če je g polinom, pa ima bistveno singularnost v neskončnosti. Ker pa ima vsaka racionalna funkcija v singularnostih kvečjemu pol, dobimo protislovje. ∎

Opomba 2. Zelo podobno vidimo, da je $\log g$ transcendenten nad $\mathbb{C}(x)$, če je g nekonstantna racionalna funkcija. V tem primeru zopet vzamemo monični polinom $p \in \mathbb{C}(x)[y]$ minimalne stopnje, za katerega naj bi veljalo

$$p(\log g) = (\log g)^n + a_{n-1}(\log g)^{n-1} + \cdots + a_0 = 0\,.$$

Če to enakost odvajamo, dobimo polinomsko enačbo

$$q(\log n) = (a'_{n-1} + n(g'/g))(\log g)^{n-1} + \cdots + (a'_0 + a_1(g'/g)) = 0\,.$$

Ker je stopnja q manjša kot stopnja p in je bil p po predpostavki minimalne stopnje, mora biti $q = 0$. Posebej mora zato veljati $a'_{n-1} = -ng'/g$ oziroma $e^{a_{n-1}} = D/g^n$. Če a_{n-1} ni konstanten, je $e^{a_{n-1}}$ transcendenten nad $\mathbb{C}(x)$ po lemi 4 in pridemo v protislovje, če pa je a_{n-1} konstanten, je tak tudi g, kar pa smo izključili.

Lema 5. *Naj bo $g \in \mathbb{C}(x)$ nekonstantna racionalna funkcija in $p \in \mathbb{C}(x)[e^g]$ poljuben polinom pozitivne stopnje n s koeficienti v $\mathbb{C}(x)$. Potem je $(p(e^g))'$ zopet polinom stopnje n v $\mathbb{C}(x)[e^g]$ in $p(e^g)$ deli $(p(e^g))'$ v $\mathbb{C}(x)[e^g]$ natanko tedaj, ko je p monom.*

Opomba 3. Preden dokažemo lemo, pokomentirajmo izraz $(p(e^g))'$. Kot smo že omenili, iz transcendentnosti e^g nad $\mathbb{C}(x)$ sledi, da sta kolobarja $\mathbb{C}(x)[e^g]$ in $\mathbb{C}(x)[y]$ izomorfna, in sicer s kanoničnim izomorfizmom, ki preslika e^g v nedoločenko y. Če je $p \in \mathbb{C}(x)[e^g]$ polinom $p(e^g) = a_n e^{ng} + a_{n-1}e^{(n-1)g} + \cdots + a_0$, kjer so $a_j \in \mathbb{C}(x)$, izraz $(p(e^g))'$ ne pomeni nič drugega kot odvajanje funkcije $p(e^{g(x)})$ po spremenljivki x. Ker je prostor $\mathbb{C}(x)[e^g]$ zaprt za odvajanje, saj je $(p(e^g))' = (a'_n + na_n g')e^{ng} + (a'_{n-1} + (n-1)a_{n-1}g')e^{(n-1)g} + \cdots + a'_0 \in \mathbb{C}(x)[e^g]$, je vprašanje, ali $p(e^g)$ deli $(p(e^g))'$, smiselno. Če s pomočjo izomorfizma $e^g \mapsto y$ razumemo polinom p kot polinom v $\mathbb{C}(x)[y]$, se ta odvod prenese na operator $\partial\colon \mathbb{C}(x)[y] \to \mathbb{C}(x)[y]$, podan z $\partial p = (a'_n + na_n g')y^n + (a'_{n-1} + (n-1)a_{n-1}g')y^{n-1} + \cdots + a'_0$. Operator ∂ seveda zadošča Leibnizovemu pravilu $\partial(pq) = p\partial q + q\partial p$. V povsem algebraičnem jeziku nam zgornja lema pove, da $p \in \mathbb{C}(x)[y]$ deli ∂p v $\mathbb{C}(x)[y]$ natanko tedaj, ko je $p = ay^n$ za neki $a \in \mathbb{C}(x)$.

Dokaz. Če je p monom, je dokaz preprost. Naj bo $f_n \in \mathbb{C}(x)$ vodilni koeficient polinoma p. Velja

$$(f_n e^{ng})' = \left(f'_n + nf_n g'\right) e^{ng}.$$

Seveda je $f'_n + nf_n g' \in \mathbb{C}(x)$ neničeln, saj bi bila drugače funkcija $f_n e^{ng}$ konstantna. Pokažimo sedaj, da $p(e^g)$ ne more deliti $(p(e^g))'$, če ima p pozitivno stopnjo in ni monom. Naj bosta $f_n e^{ng}$ in $f_m e^{mg}$ dva različna neničelna sumanda v $p(e^g)$. Ker je e^g transcendenten nad $\mathbb{C}(x)$, je $\mathbb{C}(x)[e^g]$ izomorfen $\mathbb{C}(x)[y]$, kjer je y nedoločenka. Zato je deljivost v $C(x)[e^g]$ povsem enaka deljivosti v $\mathbb{C}(x)[y]$ in mora po zgornjem izračunu veljati

$$\frac{f'_n + nf_n g'}{f_n} = \frac{f'_m + mf_m g'}{f_m}$$

oziroma

$$\left(\frac{f_n e^{ng}}{f_m e^{mg}}\right)' = 0\,.$$

Ta enakost narekuje obstoj konstante C, da velja $(f_n/f_m) = Ce^{(m-n)g}$. Ker je funkcija $e^{(m-n)g}$ transcendentna nad $\mathbb{C}(x)$, to ni mogoče. ■

Izrek 6 (Liouville, 1835). *Naj bosta $f, g \in \mathbb{C}(x)$ in g ni konstantna. Funkcija fe^g ima elementaren integral natanko tedaj, ko obstaja racionalna funkcija R, da velja*

$$f = R' + Rg'.$$

Preden se spustimo v dokaz izreka, na dveh primerih poglejmo njegovo uporabo.

Primer 4. Pokažimo, da funkcija e^{x^2} nima elementarnega integrala. Seveda od tod sledi, da Gaussov integral iz uvoda prav tako ni elementaren. Po izreku 6 ima funkcija e^{x^2} elementaren integral natanko tedaj, ko obstaja taka racionalna funkcija R, da velja

$$1 = R' + 2xR\,.$$

Takoj lahko opazimo, da R ne more biti polinom, ker bi potem desna stran enakosti morala biti polinom stopnje vsaj 1. Torej ima R vsaj en pol na kompleksni ravnini, denimo v točki a. Naj bo stopnja pola v točki a enaka k. Lokalno okoli točke a lahko R zapišemo kot

$$R(z) = \frac{\phi(z)}{(z-a)^k}\,,$$

kjer je ϕ holomorfna in neničelna v okolici točke a. Za spremembo imena spremenljivke iz x v z smo se odločili samo zaradi poudarka, da R sedaj razumemo kot holomorfno funkcijo. Odvod je tedaj v okolici a oblike

$$R'(z) = \frac{(z-a)\phi'(z) - k\phi(z)}{(z-a)^{k+1}} = \frac{\psi(z)}{(z-a)^{k+1}}\,,$$

kjer je ψ holomorfna in neničelna v okolici a. Tako ima

$$R'(z) + 2zR(z) = \frac{\psi(z) + 2z(z-a)\phi(z)}{(z-a)^{k+1}}$$

v točki a zopet pol stopnje natanko $k+1$ in ne more biti konstanta.

Primer 5. Integral $\int \frac{\mathrm{d}x}{\log x}$ lahko s substitucijo $\log x = u$ prevedemo v integral

$$\int \frac{e^u}{u}\,\mathrm{d}u\,.$$

Pokažimo, da ta integral ne more biti elementaren. Če bi bil, bi morala obstajati racionalna funkcija R, za katero bi veljalo

$$\frac{1}{u} = R' + R\,. \tag{2}$$

R očitno ne more biti polinom, zato ima R pol v vsaj eni točki $a \in \mathbb{C}$, ki naj bo pozitivne stopnje k. Desna stran (2) ima tedaj v a pol stopnje $k+1 \geq 2$. Ker ima leva stran pol samo v točki 0, ki pa je stopnje 1, (2) ne more biti izpolnjena. Z večkratno uporabo formule per partes lahko pokažemo, da je integral

$$\int \frac{e^x}{x^n}\,\mathrm{d}x$$

prav tako neelementaren za vsako naravno število n.

Dokaz (izreka 6). Pogoj je vsekakor zadosten, saj je v primeru, ko je $f = R' + Rg'$, funkcija Re^g integral fe^g. Poglejmo, da je pogoj tudi potreben. Funkcija $F(x) = f(x)e^{g(x)}$ je elementarna in se nahaja v elementarnem obsegu $\mathbb{C}(x, e^{g(x)})$. Po izreku 3 ima F elementaren integral natanko tedaj, ko jo lahko zapišemo kot

$$F(x) = \sum_{j=1}^{n} a_j \frac{\big(r_j\big(x, e^{g(x)}\big)\big)'}{r_j\big(x, e^{g(x)}\big)} + \Big(r\Big(x, e^{g(x)}\Big)\Big)',$$

kjer so $r_1, \ldots, r_n, r \in \mathbb{C}(x, y)$ racionalne funkcije v dveh spremenljivkah in $a_1, \ldots, a_n$ kompleksne konstante. Seveda lahko vsako racionalno funkcijo r_j, $j \in \{1, \ldots, n\}$ zapišemo kot $r_j = p_j/q_j$, kjer sta $p_j, q_j \in \mathbb{C}(x)[y]$. Enako lahko zapišemo $r = p/q$, $p, q \in \mathbb{C}(x)[y]$ in predpostavimo, da sta si p in q tuja v $\mathbb{C}(x)[y]$, kar pomeni, da ne obstaja polinom v $\mathbb{C}(x)[y]$ pozitivne stopnje, ki bi delil tako p kot q. Zgornja enačba tako dobi obliko

$$\begin{aligned} f(x)e^{g(x)} = &\sum_{j=1}^{n} a_j \frac{\big(p_j\big(x, e^{g(x)}\big)\big)'}{p_j\big(x, e^{g(x)}\big)} - \sum_{j=1}^{n} a_j \frac{\big(q_j\big(x, e^{g(x)}\big)\big)'}{q_j\big(x, e^{g(x)}\big)} + \\ &+ \frac{\big(p\big(x, e^{g(x)}\big)\big)' q\big(x, e^{g(x)}\big) - p\big(x, e^{g(x)}\big)\big(q\big(x, e^{g(x)}\big)\big)'}{\big(q\big(x, e^{g(x)}\big)\big)^2}\,. \end{aligned} \tag{3}$$

Vsak p_j lahko v $\mathbb{C}(x)[y]$ razcepimo na produkt $p_j = p_{j,1} \cdots p_{j,k}$, kjer so $p_{j,1}, \ldots, p_{j,k}$ bodisi nerazcepni monični polinomi (vodilni koeficient je 1) v $\mathbb{C}(x)[y]$ ali pa so elementi $\mathbb{C}(x)$. Podobno lahko razcepimo tudi polinome q_j.

Z uporabo enakosti $\frac{(st)'}{st} = \frac{s'}{s} + \frac{t'}{t}$ postane izraz (3) še preprostejši, in sicer

$$\begin{aligned} f(x)e^{g(x)} = & \sum_{j=1}^{m} a_j \frac{\big(p_j\big(x, e^{g(x)}\big)\big)'}{p_j\big(x, e^{g(x)}\big)} + \\ & + \frac{\big(p\big(x, e^{g(x)}\big)\big)' q\big(x, e^{g(x)}\big) - p\big(x, e^{g(x)}\big)\big(q\big(x, e^{g(x)}\big)\big)'}{\big(q\big(x, e^{g(x)}\big)\big)^2}, \end{aligned} \tag{4}$$

pri čemer so p_j vsi bodisi monični nerazcepni polinomi iz $\mathbb{C}(x)[y]$ bodisi elementi $\mathbb{C}(x)$. Predpostavimo lahko tudi, da so vsi p_j različni med seboj, saj v primeru ponavljanja ustrezne člene seštejemo.

Ker je po lemi 4 e^g transcendenten nad $\mathbb{C}(x)$, je preslikava, podana z $e^g \mapsto y$, izomorfizem tako med obsegoma $\mathbb{C}(x, y)$ in $\mathbb{C}(x, e^g)$ kot tudi kolobarjema $\mathbb{C}(x)[y]$ in $\mathbb{C}(x)[e^g]$. Zato je izraz (4) ekvivalenten enačbi

$$f(x)y = \sum_{j=1}^{m} a_j \frac{\partial p_j}{p_j} + \frac{q\partial p - p\partial q}{q^2} \tag{5}$$

v obsegu $\mathbb{C}(x, y)$. Operator $\partial \colon \mathbb{C}(x)[y] \to \mathbb{C}(x)[y]$ je definiran v opombi 3.

Leva stran enačbe (5) je element $\mathbb{C}(x)[y]$, medtem ko je desna stran v splošnem racionalna funkcija v dveh spremenljivkah. Poskrbeti moramo torej, da se vsi nerazcepni polinomi na desni strani krajšajo iz imenovalca. Naj bo p_k eden od nerazcepnih moničnih polinomov v (5). Po lemi 5 p_k ne more deliti ∂p_k, razen če je $p_k = y$ (potenca pri y je nujno 1, saj so faktorji nerazcepni). Predpostavimo sedaj, da p_k ni enak y. Ker so vsi p_j med seboj različni in so vsi nerazcepni, se p_k ne more pokrajšati iz imenovalca v (5) znotraj izraza $\sum_j^m a_j \partial p_j / p_j$. Da se bo pokrajšal s pomočjo izraza $(q\partial p - p\partial q)/q^2$, mora nastopiti kot nerazcepni faktor v razcepu polinoma q. Naj bo q_1 neki nerazcepen monični polinom iz $\mathbb{C}(x)[y]$, ki nastopa v razcepu polinoma q in ni enak y. Torej je $q = \tilde{q} q_1^l$ za neko pozitivno naravno potenco l in q_1 ne deli $\tilde{q}$. Dobimo

$$\frac{q\partial p - p\partial q}{q^2} = \frac{\partial p}{\tilde{q} q_1^l} - p \frac{\partial \tilde{q}}{\tilde{q}^2 q_1^l} - lp \frac{\partial q_1}{\tilde{q} q_1^{l+1}}.$$

Zopet uporabimo lemo 5 in ugotovimo, da je stopnja q_1 v imenovalcu $(q\partial p - p\partial q)/q^2$ nujno enaka $l+1 > 1$. Ker pa je stopnja p_k v imenovalcu $\sum_j^m a_j \partial p_j / p_j$ enaka 1, se ne more pokrajšati. Prav tako vidimo, da polinom q ne more imeti v razcepu nerazcepnih faktorjev različnih od y, saj bi tudi ti prispevali v imenovalec izraza (5).

Če povzamemo zgornje, dobimo, da lahko samo eden izmed p_j vsebuje tudi spremenljivko y in je enak y, preostali pa morajo biti elementi $\mathbb{C}(x)$. Prav tako mora biti q enak y^l za neki $l \in \mathbb{N} \cup \{0\}$. Zato mora biti izraz (5) oblike

$$f(x)y = \sum_j a_j \frac{\partial p_j}{p_j} + a\frac{\partial y}{y} + \frac{y^l \partial p - p\partial y^l}{y^{2l}}\,, \tag{6}$$

kjer so $p_1, \ldots, p_m \in \mathbb{C}(x)$, $p \in \mathbb{C}(x)[y]$, $l \in \mathbb{N} \cup \{0\}$ in $a, a_1, \ldots, a_m \in \mathbb{C}$. Ko nadalje polinom p razpišemo po potencah y, dobimo obliko

$$f(x)y = S(x) + ag'(x) + \sum_j \big(R_j'(x) + g'(x) j R_j(x)\big) y^j,$$

kjer indeks j teče po neki končni množici celih števil in so $R_j \in \mathbb{C}(x)$. S $S \in \mathbb{C}(x)$ smo označili vsoto $\sum_j^m a_j p_j'/p_j$. Seveda se morata izraza na levi in desni ujemati pri vsaki potenci y. Konkretno pri $j = 1$ dobimo

$$f(x) = R_1' + g' R_1\,.$$

S tem je dokaz končan. ■

Primer 6. Z majhno modifikacijo zgornjega dokaza lahko obravnavamo tudi integral funkcije $\sin x/x$. Ker velja

$$\int \frac{\sin x}{x}\,\mathrm{d}x = \int \frac{e^{ix} - e^{-ix}}{2ix}\,\mathrm{d}x = \frac{1}{2i}\int \frac{e^u - e^{-u}}{u}\,\mathrm{d}u\,,$$

bo integral elementaren natanko tedaj, ko bo elementaren integral funkcije $F(x) = (e^x - e^{-x})/x \in \mathbb{C}(x, e^x)$. Od tu naprej ravnamo povsem enako kot v dokazu izreka 6. Hitro se lahko prepričamo, da so vse manipulacije v dokazu 6 do formule (5) pravzaprav samo posledica dejstev, da operiramo znotraj obsega $\mathbb{C}(x, y) \cong \mathbb{C}(x, e^g)$ in da izraz $f(x)y \in \mathbb{C}(x, y)$ v imenovalcu nima nerazcepnih polinomov iz $\mathbb{C}(x)[y]$, različnih od y. Tako je tudi v našem primeru z $g(x) = x$ in $y = e^x$. Integral $F(x)$ bo zato elementaren natanko tedaj, ko bomo lahko zapisali

$$\frac{1}{x}y - \frac{1}{x}y^{-1} = \sum_j^m a_j \frac{\partial p_j}{p_j} + a\frac{\partial y}{y} + \frac{y^l \partial p - p\partial y^l}{y^{2l}}\,,$$

kjer so $p_1, \ldots, p_m \in \mathbb{C}(x)$, $p \in \mathbb{C}(x)[y]$, $l \in \mathbb{N} \cup \{0\}$ in $a, a_1, \ldots, a_m \in \mathbb{C}$ oziroma

$$\frac{1}{x}y - \frac{1}{x}y^{-1} = S(x) + a + \sum_j (R_j'(x) + jR_j(x))y^j\,,$$

kjer indeks j teče po neki končni množici celih števil in so $R_j \in \mathbb{C}(x)$ in $S = \sum_j^m a_j p_j'/p_j$. Posebej mora veljati $R_1' + R_1 = 1/x$. Podobno, kot smo videli v zgornjih dveh primerih, bo imela pri vsaki racionalni funkciji R_1, ki ni polinom, vsota $R_1' + R_1$ pol stopnje več kot 1 v vsaj eni točki, kar pa ne velja za $1/x$. Ker seveda R_1 ne more biti polinom, dobimo protislovje.

Zelo podobno, kot smo zgradili teorijo za funkcije oblike e^g, bi lahko poiskali kriterij za elementarnost integralov funkcij oblike $f \log g$, kjer sta f in g racionalni funkciji:

Izrek 7. *Naj bosta $f, g \in \mathbb{C}(x)$ in g ni konstantna. Funkcija $f \log g$ ima elementaren integral natanko tedaj, ko obstajata racionalna funkcija R in kompleksna konstanta C, da velja*

$$f = R' + \frac{Cg'}{g}\,.$$

Pokažimo samo, da je pogoj zadosten

$$\int f \log g \,\mathrm{d}x = \int \left(R' + C\frac{g'}{g}\right) \log g \,\mathrm{d}x = \frac{C}{2}(\log g)^2 + R \log g - \int R\frac{g'}{g}\,\mathrm{d}x\,.$$

Zadnji integral je integral racionalne funkcije in zato elementaren po Laplaceovem izreku.

Primer 7. Za vsak kompleksen $a \neq 0$ je integral

$$\int \frac{\log x}{x - a}\,\mathrm{d}x$$

neelementaren, saj bi moral biti R, ki zadošča pogoju iz izreka, oblike $R(x) = \log(x - a) - C \log x + D$ za neki konstanti $C, D \in \mathbb{C}$. Funkcija take oblike pa ni racionalna funkcija. To lahko vidimo na različne načine. Med drugim opazimo, da mora imeti $R(x)$ v a singularnost, vendar pa mora biti $R(x)\,|x - a|^{1/2}$ omejena v okolici točke a.

LITERATURA

[1] B. Conrad, *Impossibility theorems for elementary integration*, preprint.

[2] R. H. Risch, *The problem of integration in finite terms*, Trans. Amer. Math. Soc. **139** (1970), str. 605–608.

[3] M. Rosenlicht, *Integration in finite terms*, Amer. Math. Monthly **79** (1972), str. 963–972.

[4] E. A. Marchisotto in G. Zakeri, *An invitation to Integration in Finite Terms*, College Math. J. **25** (1994) 4, str. 295–308.

[5] I. Vidav, *Algebra*, Matematika – fizika 4, Mladinska knjiga, Ljubljana, 1972.

[6] I. Vidav, *Višja matematika I*, Matematika – fizika 6, Državna založba Slovenije, Ljubljana, 1987.

SODOBNE MERITVE ELEKTROMAGNETNIH LASTNOSTI PROTONOV

SIMON ŠIRCA

Fakulteta za matematiko in fiziko
Univerza v Ljubljani

PACS: 13.40.Gp, 13.60.Fz

Proton je najpreprostejše atomsko jedro, ki nam šele s pravo raziskovalno sondo, elektronskim žarkom visokih energij, razodene svoje temeljne lastnosti. V zadnjih letih so eksperimentalne raziskave statične strukture protonov in dinamičnih procesov na njih skokovito napredovale. V članku prikažem dve sodobni vrsti študija elektromagnetnih lastnosti protonov: meritve njihovih elastičnih oblikovnih faktorjev in elektromagnetnih polarizirnosti.

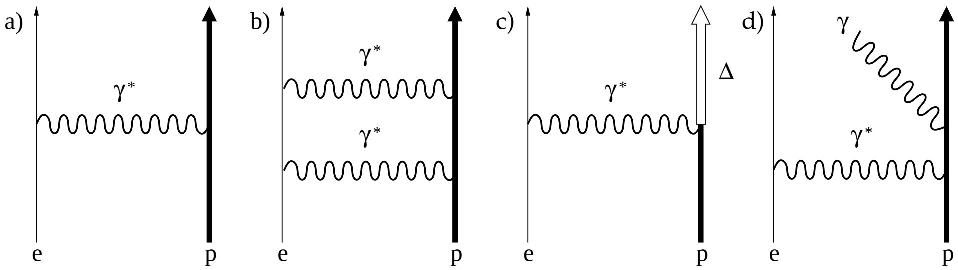
MODERN MEASUREMENTS OF ELECTRO-MAGNETIC PROPERTIES OF PROTONS

The proton is the simplest atomic nucleus that reveals its fundamental properties with a suitable probe, the high-energy electron beam. In the recent past, experimental investigations of the proton static structure and of the dynamical processes involving the proton have evolved rapidly. This paper describes two ways to study the electro-magnetic properties of protons: through measurements of elastic form-factors and electro-magnetic polarizabilities.

Rutherfordovo odkritje atomskih jeder leta 1911 je fizike naglo privedlo k spoznanju, da jedra z značilnimi prečnimi razsežnostmi nekaj femtometrov sestavljajo osnovnejši gradniki, protoni in nevtroni, s skupnim imenom nukleoni. Nukleoni nas preprosto *morajo* zanimati: predstavljajo daleč največji masni delež vidnega vesolja. Danes vemo, da imajo tudi protoni in nevtroni podstrukturo. Eksperimentalna določitev lastnosti nukleonov in umestitev izmerjenih količin v ustrezni teoretični okvir je eden osrednjih problemov ne le sodobne jedrske fizike, temveč znanosti nasploh.

Zgradbo protonov raziskujemo s sipanjem elektronov na protonskih tarčah. V energijskem območju, v katerem se gibljemo v tem članku, ima elektronska sonda pred drugimi možnimi projektili, na primer protoni, številne prednosti. Elektroni so točkasti delci in torej nimajo lastnih vzbujenih stanj. S tarčnimi jedri interagirajo zgolj z elektrošibko interakcijo, medtem ko protoni interagirajo tudi z močno (jedrsko) silo. Pri sipanju proton-proton zato vpliva obeh interakcij težko nedvoumno ločimo, saj sta projektil in tarča enaka delca.

Interakcija med elektronom in protonom poteka najpogosteje z izmenjavo enega samega virtualnega fotona (slika 1a). Virtualni foton je nosilec elektromagnetne interakcije, ki z elektrona (v kvantnomehanskem smislu, elektronskega toka) na proton (hadronski tok) prenese energijo $\omega \equiv E_\mathrm{e} - E'_\mathrm{e}$ in gibalno količino $\vec{q} \equiv \vec{p}_\mathrm{e} - \vec{p}\,'_\mathrm{e}$, kjer sta $(E_\mathrm{e}, \vec{p}_\mathrm{e})$ oziroma $(E'_\mathrm{e}, \vec{p}\,'_\mathrm{e})$ četverca gibalne količine vpadnega oziroma sipanega elektrona. Energijo vpadnih elektronov E_e določimo z meritvijo krivinskega radija njihovih poti v magnetnem polju enega ali več dipolnih magnetov. Energijo sipanih elektronov E'_e in njihov emisijski kot θ_e izmerimo z visokoločljivimi magnetnimi spektrometri. Uporabimo lahko skrajno relativistični približek $E_\mathrm{e} \approx |\vec{p}_\mathrm{e}|$ in $E'_\mathrm{e} \approx |\vec{p}\,'_\mathrm{e}|$, tako da je $|\vec{q}|^2 = E_\mathrm{e}^2 + E_\mathrm{e}'^2 - 2E_\mathrm{e}E'_\mathrm{e}\cos\theta_\mathrm{e}$. Najraje navajamo četverec prenesene gibalne količine $q^\mu \equiv (\omega, \vec{q})$ z invariantnim kvadratom $Q^2 = -q_\mu q^\mu = \vec{q}^{\,2} - \omega^2 > 0$. Pri eksperimentu lahko ω in $\vec{q}$ neodvisno spreminjamo z izbiro količin E_e, E'_e in θ_e: tako spreminjamo tudi $Q^2 \approx 4E_\mathrm{e}E'_\mathrm{e}\sin^2(\theta_\mathrm{e}/2)$. Dvofotonski proces (slika 1b) je bistveno manj verjeten, saj so tedaj namesto dveh v igri štiri elektromagnetna vozlišča: verjetnost za proces b je približno za faktor α^2 manjša od verjetnosti za proces a, kjer je $\alpha = e_0^2/4\pi\varepsilon_0\hbar c \approx 1/137$.



Slika 1. Feynmanovi grafi za sipanje elektronov na protonih: a) in b): elastično sipanje z izmenjavo enega oziroma dveh virtualnih fotonov ter c): neelastično sipanje (vzbuditev resonance Δ, prvega vzbujenega stanja nukleona). Virtualni foton γ^* iz elektrona na levi prenese na hadron na desni energijo ω in gibalno količino $\vec{q}$, ki ju lahko v eksperimentu neodvisno spreminjamo. Slika d) prikazuje virtualno comptonsko sipanje: izmenjani foton je virtualen, izsevani pa realen.

Najprimernejši so žarki elektronov z energijami od nekaj 100 MeV do nekaj GeV, saj je de Brogliejeva valovna dolžina elektronov tedaj med nekaj fm do nekaj desetink fm. Sipanje elektronov na protonih je torej neke vrste elektronski mikroskop pri zelo majhnih valovnih dolžinah. (Elektronov z ustreznimi energijami pa seveda ne moremo dobiti zgolj z elektrostatskim pospeševanjem: potrebujemo zmogljive pospeševalnike.) Sipanje je lahko elastično (slika 1a) ali neelastično (slika 1c). Izmenjata se lahko tudi dva ali več fotonov (slika 1b). Limita $Q^2 \to 0$ ustreza realnim („brezmasnim“) fotonom.

Elektromagnetni oblikovni faktorji

Naravno prvo vprašanje o ustroju protona se navezuje na prostorsko porazdelitev naboja in magnetizacije v njegovi notranjosti. Pri eksperimentu [1] merimo odvisnost elastičnega diferencialnega sipalnega preseka za elektrone na protonih,

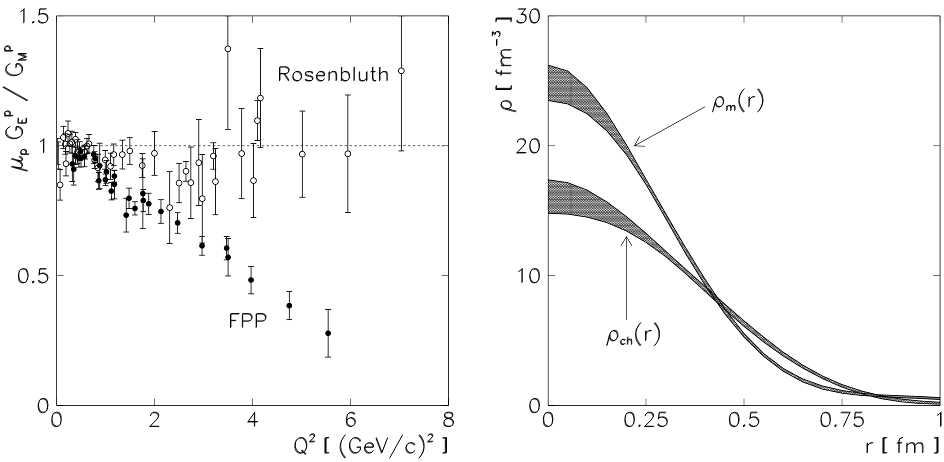
$$\frac{d\sigma}{d\Omega_e} = \frac{d\sigma_{\mathrm{Mott}}}{d\Omega_e} \frac{1}{1+\tau} \Big[\big(G_{\mathrm{E}}^{\mathrm{p}}\big)^2 + \frac{\tau}{\varepsilon} \big(G_{\mathrm{M}}^{\mathrm{p}}\big)^2 \Big] , \tag{1}$$

kjer je ε stopnja polarizacije virtualnega fotona in $\tau = Q^2/4M^2$ (mirovna masa protona je $M \approx 938\,\mathrm{MeV/c^2}$). V tem izrazu je neposredno izražen razcep sipalnega preseka, ki sledi iz približka izmenjave enega fotona: kinematika elektrona pri danem Q^2 določa kotno odvisnost in kinematične faktorje ε in τ, medtem ko je fizikalna informacija o hadronskem delu procesa shranjena v količinah $G_{\mathrm{E}}^{\mathrm{p}}(Q^2)$ in $G_{\mathrm{M}}^{\mathrm{p}}(Q^2)$. Mottov sipalni presek izraža odvisnost od sipalnega kota θ_e. V vodilnem redu vsebuje značilni potek $\sin^{-4}(\theta_e/2)$ klasičnega Rutherfordovega preseka. Upoštevati moramo tudi spinsko vrtilno količino elektrona in protona: oba sta fermiona s spinom 1/2 in ju moramo obravnavati v okviru relativistične (Diracove) enačbe, od koder dobimo dodatno kotno odvisnost $\cos^2(\theta_e/2)$. Nazadnje vključimo še majhen popravek zaradi končne mase (odrivne energije) protona.

Mottov presek velja za sipanje elektronov na točkastih protonih. Količini $G_{\mathrm{E}}^{\mathrm{p}}(Q^2)$ in $G_{\mathrm{M}}^{\mathrm{p}}(Q^2)$ pa opisujeta tisto najzanimivejše: popravek preseka zaradi končne razsežnosti in elektromagnetne podstrukture protona. Imenujemo ju električni oziroma magnetni *oblikovni faktor* in ju je mogoče natančno izmeriti. Do tega močnega pretresa v zgodovini hadronske fizike je vodila trnova pot. Ker so protoni fermioni, bi jim po Diracovi enačbi pripadal magnetni moment $+1\mu_{\mathrm{B}} = e_0\hbar/2M$ (jedrski Bohrov magneton), vendar so poskusi pokazali, da imajo protoni v resnici magnetni moment $\mu_{\mathrm{p}} \approx +2{,}79\,\mu_{\mathrm{B}}$. Ker se je izmerjeni magnetni moment tako močno razlikoval od pričakovane Diracove vrednosti, so ga imenovali „anomalni“. Kasneje so domnevali, da so protoni kljub tej razliki točkasti. Šele Hofstadterjeve legendarne meritve [2], za katere je bila pozneje podeljena Nobelova nagrada, so pokazale, da je pravilna le tretja pot: protoni imajo anomalni magnetni moment *ravno zato, ker imajo podstrukturo*.

Z realnimi fotoni ($Q^2 = 0$) otipamo le celotni naboj protona $+1\,e_0$ in izmerimo le statični magnetni moment protona μ_{p}. Zato sta oblikovna faktorja normirana na $G_{\mathrm{E}}^{\mathrm{p}}(0) = 1$ oziroma $G_{\mathrm{M}}^{\mathrm{p}}(0) = \mu_{\mathrm{p}}$. Oblikovna faktorja pri $Q^2 > 0$ določamo z elastičnim sipanjem elektronov. Iz meritve ju izločimo tako, da presek (1) izmerimo pri različnih energijah vpadnih elektronov E_e in različnih sipalnih kotih θ_e ob konstantnem Q^2. S tem spreminjamo faktor τ/ε pred $\big(G_{\mathrm{M}}^{\mathrm{p}}\big)^2$, medtem ko ostane faktor pred $\big(G_{\mathrm{E}}^{\mathrm{p}}\big)^2$ nespremenjen. To je tako imenovana Rosenbluthova ali longitudinalno-transverzalna separacija sipalnega preseka, saj z njo iz preseka posebej izluščimo električni (nabojni, longitudinalni) in magnetni (transverzalni) prispevek.

Slika 2 (levo, prazni simboli) prikazuje rezultate meritev po Rosenbluthovi metodi, izražene z razmerjem oblikovnih faktorjev. Električni člen prevladuje pri sipanju pri majhnih kotih, magnetni pa pri velikih kotih, kjer že močno pade Mottov presek: napaka se zato pri visokih vrednostih Q^2 močno poveča in pri meritvah v območju visokih Q^2 so razmerja oblikovnih faktorjev v okviru merskih napak konsistentna z ena.



Slika 2. Levo: razmerje električnega in magnetnega oblikovnega faktorja protona, izmerjeno z elastičnim sipanjem po Rosenbluthovi metodi (enačba (1)) ali s protonsko polarimetrijo (enačba (3)), v odvisnosti od prenosa gibalne količine; desno: radialni gostoti naboja in magnetizacije, določeni iz izmerjenih oblikovnih faktorjev (črni pas označuje razpon vrednosti zaradi merskih napak).

Oblikovna faktorja opisujeta od Q^2 odvisno sklopitev električnega in magnetnega dela hadronskega toka z virtualnim fotonom. S spreminjanjem prenosa gibalne količine $Q^2 \neq 0$ (v prostoru gibalnih količin) „prečesavamo“ ustrezni porazdelitvi naboja in magnetizacije v konfiguracijskem prostoru. V primeru porazdelitve naboja $\rho(\vec{r})$ v brezspinskem jedru povezavo med prostoroma v nerelativistični sliki predstavlja matrični element električnega potenciala $V(\vec{r}) \propto \rho(\vec{r})$ med stanjema vpadnega in sipanega elektrona

$$\mathcal{M} \propto \int \psi^*_{e'}(\vec{r}) V(\vec{r}) \psi_e(\vec{r})\, d^3\vec{r} \approx \int e^{-i\vec{p}'_e \vec{r}} V(\vec{r}) e^{i\vec{p}_e \vec{r}} d^3\vec{r} = \int e^{i\vec{q}\,\vec{r}} V(\vec{r})\, d^3\vec{r}.$$

Nabojni oblikovni faktor $F(\vec{q}^{\,2})$ definiramo s Fourierovo transformacijo porazdelitve naboja

$$F(\vec{q}^{\,2}) = \int \rho(\vec{r}) e^{i\vec{q}\,\vec{r}} d^3\vec{r},$$

merjeni presek pa je sorazmeren z $|\mathcal{M}|^2$, torej tudi z $\left|F(\vec{q}^{\,2})\right|^2$. Porazdelitev naboja nato dobimo iz izmerjenega oblikovnega faktorja z inverzno Fourierovo transformacijo. Proton ni brezspinsko jedro, zato ima dva oblikovna faktorja (električnega $G_{\mathrm{E}}^{\mathrm{p}}$ in magnetnega $G_{\mathrm{M}}^{\mathrm{p}}$) in za elastično sipanje pri energijah nekaj GeV ne velja nerelativistična slika, zato Fourierovo transformacijo nadomesti splošnejša transformacija

$$\rho_{\mathrm{ch}}(r) = \frac{2}{\pi}\int\limits_0^\infty \mathrm{d}k\, k^2 j_0(kr) G_{\mathrm{E}}^{\mathrm{p}}(Q^2)(1+\tau)^{\lambda_{\mathrm{E}}},$$

$$\mu_{\mathrm{p}}\, \rho_{\mathrm{m}}(r) = \frac{2}{\pi}\int\limits_0^\infty \mathrm{d}k\, k^2 j_0(kr) G_{\mathrm{M}}^{\mathrm{p}}(Q^2)(1+\tau)^{\lambda_{\mathrm{M}}},$$

kjer je $k^2 = Q^2/(1+\tau)$. (Majhna popravka $(1+\tau)^{\lambda_{\mathrm{E}}}$ in $(1+\tau)^{\lambda_{\mathrm{M}}}$ k običajni obliki sta potrebna, ker je nerelativistična interpretacija oblikovnih faktorjev modelsko odvisna. To pomeni, da moramo pri nerelativistični redukciji relativističnih matričnih elementov napraviti določene privzetke o obliki interakcije, ki pa je vnaprej ne poznamo. Parametra λ_{E} in λ_{M} dovolj dobro utelešata vse take privzetke: običajno vzamemo $\lambda_{\mathrm{E}} = \lambda_{\mathrm{M}} = 2$, vendar tudi druge izbire rezultate le malo spremenijo [3].) Šele s to preslikavo torej dobimo pravilni in od modelskih približkov skoraj neodvisni radialni gostoti naboja in magnetizacije v protonu.

Strmina električnega oblikovnega faktorja pri majhnih Q^2 je povezana s pričakovano vrednostjo kvadrata radija protona. Inverzno Fourierovo transformacijo razvijemo po potencah Q^2, $G_{\mathrm{E}}^{\mathrm{p}}(Q^2) = 1 - \frac{1}{6}Q^2\langle r_{\mathrm{E}}^2\rangle_{\mathrm{p}} + \cdots$, od koder sledi

$$\langle r_{\mathrm{E}}^2\rangle_{\mathrm{p}} = -6\,\frac{\mathrm{d}G_{\mathrm{E}}^{\mathrm{p}}}{\mathrm{d}Q^2}\bigg|_{Q^2\to 0} \approx 0{,}69\ \mathrm{fm}^2. \tag{2}$$

Značilna razsežnost porazdelitve naboja protona oziroma njegov „polmer“ je torej $\sqrt{\langle r_{\mathrm{E}}^2\rangle_{\mathrm{p}}} \approx 0{,}83$ fm.

Do nedavnega je bil Rosenbluthov način meritve z ločitvijo longitudinalnih in transverzalnih parcialnih presekov edina metoda za določitev obeh oblikovnih faktorjev hkrati. Glavna hiba te metode so velike sistematične napake, ki izvirajo predvsem iz močne kotne odvisnosti sipalnega preseka. Na izide starejših poskusov je vplivalo tudi pomanjkljivo poznavanje sprejemljivosti detektorjev, ploskovne gostote tarčnih jeder in toka vpadnih elektronov.

Oblikovna faktorja je mogoče določiti tudi z novim pristopom, pri katerem izkoristimo spinske prostostne stopnje [4]. Za meritev potrebujemo polariziran elektronski žarek, izmeriti pa moramo tudi stopnjo polarizacije iz elastičnega sipanja odrinjenih protonov. Zaradi velike instrumentalne zahtevnosti so tovrstne meritve izvedljive šele zadnjih nekaj let. Ključna je meritev polarizacije protonov: zanjo potrebujemo poseben polarimeter, ki ga vgradimo v visokoločljiv magnetni spektrometer, s katerim sicer lahko analiziramo tudi nepolarizirane delce. S polarimetrom izmerimo komponenti polarizacije P_l in P_t (oziroma ustrezna enotska vektorja) v ravnini, ki jo določata gibalni količini vpadnih in sipanih elektronov. Natančna izpeljava daleč presega namen tega članka, vendar je na koncu razmerje oblikovnih faktorjev mogoče zapisati v preprosti obliki

$$\frac{\mu_\mathrm{p} G_\mathrm{E}^\mathrm{p}(Q^2)}{G_\mathrm{M}^\mathrm{p}(Q^2)} = -\frac{P_\mathrm{t}}{P_\mathrm{l}} \frac{E_\mathrm{e} + E_\mathrm{e}'}{2M} \tan\frac{\theta_\mathrm{e}}{2} \,. \tag{3}$$

Razmerje $G_\mathrm{E}^\mathrm{p}(Q^2)/G_\mathrm{M}^\mathrm{p}(Q^2)$ torej lahko z zelo majhno sistematsko napako določimo z meritvijo razmerja $P_\mathrm{t}/P_\mathrm{l}$. Napake meritev kinematičnih količin na desni strani enačbe so zanemarljive. Slika 2 (levo, polni simboli) prikazuje razmerje oblikovnih faktorjev po polarizacijski metodi. Na prvi pogled je razviden približno linearen padec razmerja v odvisnosti od Q^2, v močnem nasprotju z Rosenbluthovo metodo. Kaj to pomeni? Medtem ko z Rosenbluthovo analizo (1) v okviru merskih napak dobimo skoraj enako radialno odvisnost $\rho_\mathrm{ch}(r)$ in $\rho_\mathrm{m}(r)$, z metodo (3) ugotovimo, da je magnetizacija v protonu porazdeljena po manjšem območju (znotraj manjšega volumna) kot njegov naboj. (Slika 2 (desno) prikazuje radialni gostoti naboja in magnetizacije v protonu, ki ju dobimo po polarizacijski metodi.)

Kateri pristop je pravi? Sodobni teoretični izračuni kažejo [5], da so pri elastičnem sipanju pri visokih prenosih gibalne količine zelo pomembni tudi prispevki višjih redov elektromagnetne interakcije, pri kateri se izmenja več fotonov (slika 1b). Razlika med metodama še ni dokončno pojasnjena, vendar je učinek izmenjave dveh ali več fotonov na razmerje oblikovnih faktorjev bistveno močnejši pri Rosenbluthovi metodi (četudi se sipalni presek spremeni le malo) kot pri polarizacijski metodi. Polarizacijska metoda torej izmeri „bolj prava“ oblikovna faktorja protona, ki ustrezata porazdelitvama naboja in magnetizacije kot na sliki 2 (desno).

Polarizirnost

Elektroni v elektronskem oblaku električno nevtralnega atoma, ki je izpostavljen zunanjemu električnemu polju, čutijo električno silo $\vec{F} = m\omega^2\vec{x} = e_0\vec{E}$, kjer $m\omega^2$ karakterizira konstanto „vzmeti“, s katero je posamezni elektron pripet k jedru. Zunanje polje deformira elektronski oblak in v atomu inducira električni dipolni moment $\vec{p}_\mathrm{e} = e_0\vec{x} = (e_0^2/m\omega^2)\,\vec{E} \equiv \alpha\vec{E}$. Sorazmernostni koeficient med induciranim momentom in zunanjim električnim poljem je električna polarizirnost. Velikostni red električne polarizirnosti atoma je $\alpha \approx 10^{-28}$ m^3, kar je enakega velikostnega reda kot prostornina, ki jo zavzema elektronski oblak. Premik težišča elektronskega oblaka glede na jedro pri $E = 10\,\mathrm{kV/cm}$ je $x \approx 10^{-6}$ nm, kar je še merljivo s sodobnimi interferometričnimi metodami [7].

Električno (in magnetno) polarizirnost protonov, pri katerih pričakujemo relativni premik pozitivnih in negativnih nabojev (oziroma delno reorientacijo magnetnih momentov) v protonu, lahko definiramo po analogiji s polarizirnostjo atomov. Zunanje električno polje $\vec{E}$ in magnetno polje $\vec{H}$ inducirata v protonu električni in magnetni dipolni moment $\vec{p}_\mathrm{e} = \alpha_\mathrm{p}\vec{E}$ oziroma $\vec{p}_\mathrm{m} = \beta_\mathrm{p}\vec{H}$. Sorazmernostni konstanti imenujemo statična električna oziroma magnetna polarizirnost protona.

Za proton z radijem nekaj desetink fm bi po primerjavi z atomskim svetom zgolj z dimenzijskimi argumenti pričakovali $\alpha_\mathrm{p} \approx 10^{-18}\,\alpha \approx 10^{-3}$ fm^3. Zanesljiveje pa električno in magnetno polarizirnost za protone ocenimo v drugem redu teorije motenj. (Prvi red k polarizirnosti ne prispeva: matrični elementi vektorskega operatorja $\vec{p}_\mathrm{e}$ med stanjema z enako parnostjo so enaki nič.) Sprememba energije protona v statičnem električnem polju, ki naj kaže v smeri osi z, je enaka

$$\Delta W_E = -\sum_{\mathrm{X}\neq\mathrm{p}} \frac{|\langle \mathrm{p}| - \widehat{p}_{\mathrm{e}z}E|\mathrm{X}\rangle|^2}{M_\mathrm{X} - M_\mathrm{p}} = -\int_0^E \vec{p}_\mathrm{e}\,\mathrm{d}\vec{E} = -\frac{1}{2}\,\alpha_\mathrm{p}E^2 .$$

V enačbo vrinemo kompletni ortonormirani sistem $\sum_\mathrm{X} |\mathrm{X}\rangle\langle\mathrm{X}| = 1$. Zunanje električno polje lahko vzbudi le električne dipolne prehode iz osnovnega stanja $|\mathrm{p}\rangle$ (protona) s pozitivno parnostjo v stanja $|\mathrm{X}\rangle$ z negativno parnostjo. Nekaj takšnih vzbujenih stanj (nukleonskih resonanc) leži od 600 do 700 MeV nad osnovnim stanjem, na primer N(1520), N(1535) in N(1650) [6]. Za oceno zato zadošča, če za energijski imenovalec vstavimo kar povprečno vrednost $\Delta M_\mathrm{Xp} \equiv \overline{M_\mathrm{X} - M_\mathrm{p}} \approx 700\,\mathrm{MeV}$. Dobimo

$$\alpha_\mathrm{p} \approx \frac{2\sum_\mathrm{X}\langle\mathrm{p}|\,\widehat{p}_{\mathrm{e}z}\,|\mathrm{X}\rangle\,\langle\mathrm{X}|\,\widehat{p}_{\mathrm{e}z}\,|\mathrm{p}\rangle}{\Delta M_\mathrm{Xp}} = \frac{2}{\Delta M_\mathrm{XN}}\langle\mathrm{p}|\,e_0^2\,\tfrac{1}{3}\,r^2\,|\mathrm{p}\rangle = \frac{2e_0^2}{3\Delta M_\mathrm{Xp}}\left\langle r_\mathrm{E}^2\right\rangle_\mathrm{p} .$$

S povprečnim kvadratom električnega radija protona $\left\langle r_\mathrm{E}^2\right\rangle_\mathrm{p} \approx 0{,}69$ fm^2 (enačba (2)) dobimo oceno $\alpha_\mathrm{p} \approx 10 \cdot 10^{-4}\,\mathrm{fm}^3$. To je povsem v skladu z volumskim razmislekom pri atomski polarizirnosti. Po enakem kopitu dobimo tudi oceno za magnetno polarizirnost $\beta_\mathrm{p} \approx 2 \cdot 10^{-4}\,\mathrm{fm}^3$.

Očitno je pri laboratorijsko dosegljivih električnih poljskih jakostih ustrezni premik težišča nabojev nemerljivo majhen (oziroma proton izjemno „neraztegljiv“). Vendar lahko polarizirnost protona vseeno izmerimo, in sicer s comptonskim sipanjem fotonov na protonih. Primerne fotone dobimo s pomočjo istih elektronskih žarkov, ki jih potrebujemo pri študiju oblikovnih faktorjev: v žarek postavimo ploščico trdožive snovi, v kateri elektroni zavorno sevajo. V skrajnem primeru izsevani foton odnese vso energijo elektrona. Tako lahko dobimo realne fotone z energijami nekaj 100 MeV do nekaj GeV.

V curku fotonov, usmerjenem na protonsko tarčo, se polji $\vec{E}$ in $\vec{H}$ spreminjata s časom, torej oscilirata tudi inducirana dipolna momenta $\vec{p}_{\mathrm{e}}$ in $\vec{p}_{\mathrm{m}}$, zato v klasični sliki poleg vpadnega valovanja (z energijo $\hbar\omega$) dobimo tudi sekundarno električno in magnetno dipolno sevanje — sipane fotone (z energijo $\hbar\omega'$). Te fotone s scintilacijskimi detektorji detektiramo in analiziramo po energijah in sipalnih kotih θ. Diferencialni sipalni presek za comptonsko sipanje fotonov na protonih z energijami fotonov, mnogo manjšimi od mirovne mase pionov, ima obliko [8]

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}=\frac{\mathrm{d}\sigma_{\mathrm{Powell}}}{\mathrm{d}\Omega}-\omega\omega'\frac{e_0^2}{M}\left(\frac{\omega'}{\omega}\right)^2\left\{\frac{\alpha_{\mathrm{p}}+\beta_{\mathrm{p}}}{2}(1+\cos\theta)^2+\frac{\alpha_{\mathrm{p}}-\beta_{\mathrm{p}}}{2}(1-\cos\theta)^2\right\}, \quad (4)$$

kjer sta ω in ω' povezani s Comptonovo relacijo $\omega' = \omega\big[1 + \frac{\omega}{M}(1 - \cos\theta)\big]^{-1}$. Od tod je igra na moč podobna tisti pri določevanju oblikovnih faktorjev: del preseka nosi odvisnost od znanih, z merskimi napravami neposredno izmerljivih kinematičnih količin ω, ω' in θ, medtem ko je fizikalno najzanimivejša vsebina skrita v strukturnih parametrih α_{p} in β_{p}, ki opisujeta odziv protona na elektromagnetno motnjo. Prvi (Powellov) člen v izrazu (4) ustreza comptonskemu sipanju na točkastem protonu (brez notranje strukture), vendar z anomalnim magnetnim momentom ($\mu_{\mathrm{p}} - 1$) [9]. (Vodilni člen tega izraza je Klein-Nishinova formula za sipanje nepolariziranih fotonov na delcih brez strukture z maso M in nabojem e_0 [10].) V limiti $\omega \to 0$ preide člen v dobro znano Thomsonovo formulo

$$\lim_{\omega\to 0}\frac{\mathrm{d}\sigma_{\mathrm{Powell}}}{\mathrm{d}\Omega} = \frac{r_0^2}{2}\left(1+\cos^2\theta\right),$$

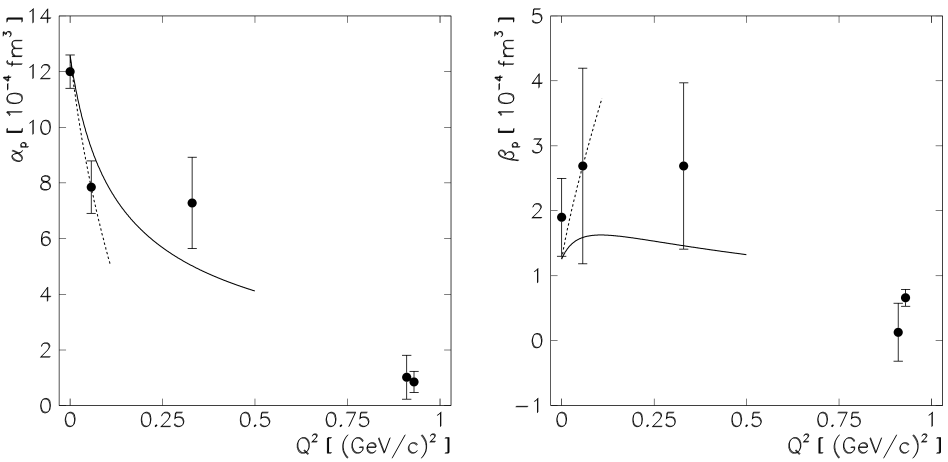
kjer je r_0 klasični radij delca z maso M. (Thomsonova limita sipalnega preseka sledi iz klasične elektrodinamike in ne vsebuje kvantnih efektov.) Powellov presek je eksaktno izračunljiv, celotni diferencialni sipalni presek (4) pa merimo pri poskusu, tako da sta parametra α_{p} in β_{p} edini neznanki.

Iz enačbe (4) je razvidno, da je presek za sipanje v $\theta = 0°$ občutljiv le na vsoto $\alpha_{\mathrm{p}} + \beta_{\mathrm{p}}$, presek za sipanje v $\theta = 180°$ pa le na razliko $\alpha_{\mathrm{p}} - \beta_{\mathrm{p}}$. Prav to kotno odvisnost izkoriščajo pri meritvah comptonskega sipanja fotonov na protonih. (In prav zato sta merski napaki za α_{p} in β_{p} antikorelirani.) Iz meritve pri $\theta = 90°$ lahko osamijo α_{p}, medtem ko polarizirnosti β_{p} ne moremo izmeriti ločeno. Povprečji svetovnih meritev [8] sta $\alpha_{\mathrm{p}} = (12{,}0 \pm 0{,}6) \cdot 10^{-4}$ fm^3 oziroma $\beta_{\mathrm{p}} = (1{,}9 \mp 0{,}6) \cdot 10^{-4}$ fm^3. Vsota električne in magnetne polarizirnosti je omejena z Baldinovim vsotnim pravilom

$$\alpha_\mathrm{p} + \beta_\mathrm{p} = \frac{1}{2\pi^2} \int\limits_\mathrm{prag}^{\infty} \frac{\sigma_\mathrm{tot}(\omega)}{\omega^2}\,\mathrm{d}\omega = (13{,}9 \pm 0{,}3) \cdot 10^{-4}\ \mathrm{fm}^3.$$

To vsotno pravilo je izpeljano iz zelo strogih teoretičnih predpostavk in poveže vsoto polarizirnosti z integralom totalnega fotoabsorpcijskega preseka, uteženega z $1/\omega^2$, ter ga je mogoče preveriti neodvisno.

V zadnjih nekaj letih so napravile meritve polarizirnosti še velikanski korak naprej. Namesto realnih fotonov uporabimo virtualne: proton pri elektronskem sipanju najprej absorbira virtualni foton, nato pa izseva realni foton (virtualno comptonsko sipanje, slika 1d). Detektiramo odrinjeni proton. Podobno kot z elektromagnetnimi oblikovnimi faktorji z meritvijo presekov za virtualno comptonsko sipanje in določitvijo ustreznih *posplošenih polarizirnosti* ugotavljamo, kakšno prostorsko porazdelitev imata električna in magnetna polarizacija v protonu. S posplošenimi polarizirnostmi torej izrazimo, na kakšen način in v kolikšni meri se sestavni deli protona, kvarki in pioni, preurejajo v zunanjem elektromagnetnem polju. Povsem analogno s kvadratom radija naboja (enačba (2)) definiramo tudi pričakovano vrednost kvadrata električnega polarizirnostnega radija s strmino posplošene električne polarizirnosti $\langle r_\alpha^2 \rangle_\mathrm{p} \propto -6\left(\mathrm{d}\alpha_\mathrm{p}/\mathrm{d}Q^2\right)$ v limiti $Q^2 \to 0$.



Slika 3. Posplošena električna (levo) in magnetna (desno) polarizirnost protona v odvisnosti od prenosa gibalne količine pri virtualnem comptonskem sipanju. Polna krivulja predstavlja teoretični izračun polarizirnosti, ki pri $Q^2 > 0{,}5$ (GeV/c)^2 postane nezanesljiv [12], črtkana krivulja pa pričakovano obnašanje α_p in β_p pri zelo nizkih Q^2.

Meritve posplošenih polarizirnosti z virtualnim comptonskim sipanjem so na frontni črti eksperimentalne hadronske fizike. Iz centrov MIT-Bates, MAMI in Jefferson Laboratory [11] imamo zato za zdaj na voljo le peščico rezultatov, iz katerih pa že lahko v grobem razberemo dve izstopajoči potezi (slika 3). Relativno velika pričakovana vrednost $\langle r_\alpha^2 \rangle_\mathrm{p}$ nam pove, da v posplošeni električni polarizirnosti protona $\alpha_\mathrm{p}(Q^2)$ prevladujejo mezonski efekti: kvazistatično električno polje virtualnih fotonov polarizira proton tako, da premakne težišče pozitivno nabitega pionskega oblaka, ki obdaja kvarkovsko sredico. Po drugi strani je magnetna polarizirnost $\beta_\mathrm{p}(Q^2)$ nekajkrat manjša od električne in se v odvisnosti od prenosa gibalne količine le počasi spreminja: to razložimo s paramagnetnim prispevkom kvarkovske sredice, ki je skoraj nasprotno enak diamagnetnemu prispevku pionskega oblaka.

LITERATURA

[1] I. A. Qattan et al., *Precision Rosenbluth Measurement of the Proton Elastic Form Factors*, Phys. Rev. Lett. **94** (2005), 142301.

[2] R. Hofstadter, F. Bumiller in M. R. Yearian, *Electromagnetic Structure of the Proton and Neutron* , Rev. Mod. Phys. **30** (1958) 2, str. 482–497.

[3] J. J. Kelly, *Nucleon Charge and Magnetization Densities from Sachs Form Factors*, Phys. Rev. C **66** (2002), 065203.

[4] O. Gayou et al., *Measurement of* G_{E_p}/G_{M_p} *in* $\vec{e}p \to e\vec{p}$ *to* $Q^2 = 5.6$ *GeV*^{2}, Phys. Rev. Lett. **88** (2002) 9, 092301.

[5] D. Day, *Nucleon elastic form factors – Current status of the experimental effort*, Eur. Phys. J. A **31** (2007) 4, str. 560–565.

[6] W.-M. Yao et al. (Particle Data Group), *Review of particle physics*, J. Phys. G: Nucl. Part. Phys. **33** (2006), str. 1–1232.

[7] M. D. Bonin in V. Kresin, *Electric-dipole polarizabilities of atoms, molecules and clusters*, World Scientific, Singapore, 1997.

[8] M. Schumacher, *Polarizability of the Nucleon and Compton Scattering*, Prog. Part. Nucl. Phys. **55** (2005), str. 567–646.

[9] J. L. Powell, *Note on the Bremsstrahlung Produced by Protons*, Phys. Rev. **75** (1949) 1, str. 32–34.

[10] V. B. Berestetskii, E. M. Lifshits in L. P. Pitaevskii, *Quantum electrodynamics*, Pergamon Press, Oxford, 1980.

[11] P. Bourgeois et al. (OOPS Collaboration), *Measurements of the Generalized Electric and Magnetic Polarizabilities of the Proton at Low* Q^2 *Using the Virtual-Compton-Scattering Reaction*, Phys. Rev. Lett. **97** (2006), 212001; J. Roche et al., *First Determination of Generalized Polarizabilities of the Proton by a Virtual Compton Scattering Experiment*, Phys. Rev. Lett. **85** (2000) 4, str. 708–711; G. Laveissière et al. (Hall A Collaboration), *Measurement of the Generalized Polarizabilities of the Proton in Virtual Compton Scattering at* $Q^2 = 0.92$ *and 1.76 GeV*^{2}, Phys. Rev. Lett. **93** (2004) 12, 122001.

[12] T. R. Hemmert et al., *Generalized Polarizabilities and the Chiral Structure of the Nucleon*, Phys. Rev. Lett. **79** (1997) 1, str. 22–25; T. R. Hemmert et al., *Generalized polarizabilities of the nucleon in chiral effective theories*, Phys. Rev. D **62** (2000), 014013.

O REALNIH ŠTEVILIH

PETER ŠEMRL

Fakulteta za matematiko in fiziko
Univerza v Ljubljani

Math. Subj. Class. (2000): 12D99

Večina srednješolcev realna števila intuitivno dobro razume. Matematično korektno vpeljati realna števila pa še zdaleč ni enostavno. Kaj lahko o tem povemo dijakom v okviru matematičnih krožkov?

ON REAL NUMBERS

Most of secondary school pupils have a reasonable intuitive understanding of real numbers. However, introducing reals in a mathematically correct way is not an easy task. Can we discuss these problems at the secondary school level?

V zadnjih desetletjih poizkušajo šole pri nas in v primerljivih državah postati učencem bolj prijazne. Tako je tudi prav. Vsi mladostniki naj bi opravili srednješolsko izobraževanje, večina pa naj bi si pridobila tudi fakultetno izobrazbo. Tudi prizadevanja v tej smeri lahko samo pozdravimo.

Bojim pa se, da ob vseh teh hvalevrednih namenih pozabljamo na najbolj nadarjene učence in dijake. Ker dandanes mladostniki niso niti bolj pametni niti bolj delavni kot nekoč, je mogoče višjo izobrazbeno raven doseči le, če ob vseh drugih ukrepih in spremembah znižamo minimalne zahteve, ki jih morajo dijaki izpolniti za dokončanje srednje šole. Ob tako znižanih kriterijih pa nadarjeni močno izstopajo iz povprečja, ne da bi se za to kaj posebej trudili. In zato ne razvijejo svojih sposobnosti toliko, kot bi jih v bolj spodbudnem okolju.

Poleg tega na vseh nivojih izobraževanja opažamo težnjo po zmanjševanju obsega predmetov, ki so za učence in dijake težki. To je seveda najlažja pot k „zviševanju“ izobrazbene strukture prebivalstva.

Učitelji v takem okolju so seveda postavljeni pred povsem druge izzive kot učitelji v deželah s poudarjeno tekmovalnim šolskim sistemom (npr. v nekaterih azijskih državah). Dober in prizadeven učitelj se zaveda slabosti in prednosti posameznih izobraževalnih postopkov. Tako ga veselijo marsikateri pozitivni premiki v našem šolstvu, hkrati pa poizkuša zmanjševati negativne učinke sprememb, ki jih ni malo, saj so naše šole ves čas v procesu prenove ali pa reforme, kakorkoli se že pač temu v določenem času reče. In le kdo verjame, da vse vodijo v pravo smer?

Iz vsega povedanega ste seveda uganili, da imam v mislih dodatno spodbujanje najbolj nadarjenih učencev. Številni učitelji matematike in fizike, ki vodijo matematične in fizikalne krožke, opravijo delo neprecenljive vrednosti. Precej časa je namenjenega pripravi za tekmovanja iz matematike, fizike, logike in računalništva. Mnogi mentorji pa v okviru krožkov pripravijo tudi kakšno predavanje, primerno za najboljše dijake. Morda je pri pripravi takega predavanja najtežji prav prvi korak, to je izbira teme. V tem prispevku nameravam predstaviti temo, ki je primerna za matematične krožke na gimnazijah. V uredništvu si želimo več takih prispevkov, ki bi prispevali k bogatitvi vsebin pri mentorskem delu z najbolj nadarjenimi dijaki.

Kaj želimo doseči z izborom teme? V srednji šoli sem posamezna poglavja pri predmetu matematika dojemal kot zaključene zgodbe. Povedali so mi, kaj je odvod funkcije, naučili smo se ga računati in spoznali njegove uporabe. Le kaj bi tukaj lahko še kdo dodal? In podobno je bilo z integralom: nedoločeni integral, določeni integral, računanje in uporaba. Mislil sem si, da bom na študiju matematike moral računati težje integrale, o samih integralih pa se ne da povedati kaj drugega od tistega, kar sem že slišal.

Ne vidim nobenega razloga, da bi boljšim dijakom prikrivali resnico. Na kratko jim lahko povemo, da so vse to šele začetki vznemirljivo zanimivih zgodb. Na primer, povemo jim, da imamo vsepovsod opraviti s količinami, ki so odvisne od več spremenljivk. Cena izdelka je odvisna od cene vloženega dela, cene surovin, davkov, stroškov distribucije in marketinga, razmerja med ponudbo in povpraševanjem ... Zato v matematiki poleg srednješolcem znanih „običajnih“ funkcij (mišljene so seveda realne funkcije ene spremenljivke) obravnavamo tudi funkcije več spremenljivk. Potem pa postane študij odvodov, naraščanja, padanja, konveksnosti ... še veliko bolj zanimiv. Določeni integral definiramo z mislijo na računanje ploščin ravninskih likov. Kaj pa računanje volumnov in površin prostorskih teles? In potem lahko omenimo, da taka vprašanja vodijo od integralov, označenih z eno „integralsko kljuko“, k takim z dvema ali celo tremi.

V avgustu 2006 sem bil povabljen, da pripravim predavanje za srednješolce v okviru poletnega srečanja MARS 2006 v Kopru. Ob izbiranju teme sem se poleg zgoraj zapisanih misli spomnil še nekaterih pogovorov z znanci ob začetku mojega raziskovalnega dela v matematiki. Bilo jim je jasno, da se znanosti, kot medicina, biologija ali kemija, ves čas razvijajo. Nemalo pa jih je bilo presenečenih nad dejstvom, da se to dogaja tudi s tako „staro znanostjo“, kot je matematika. Da je v matematiki mogoče odkriti še kaj novega?

Izbiral sem torej tematiko, ki bi bila zanimiva, dostopna nadarjenim srednješolcem in bi med drugim pokazala, da tudi na prvi pogled dokončane zgodbe v matematiki skrivajo v sebi še mnogo zanimivega. Poučen primer je pojem realnih števil. Na začetku predavanja smo skupaj z dijaki „ugotovili“, da o številih vemo tako rekoč vse. Vsi „vemo“, kaj so števila in kako z njimi računamo. Le kaj bi tu lahko še dodali?

Potem pa smo opravili hiter sprehod skozi številske množice. V množici naravnih števil nimamo problemov s seštevanjem in množenjem. S primeri pokažemo, da se zatakne pri odštevanju, in se nato spomnimo, da ta problem odpravimo z vpeljavo celih števil. Problem deljenja pa odpravimo z uvedbo racionalnih števil. Je racionalno število isto kot ulomek? Ne, saj dva različna ulomka lahko predstavljata isto racionalno število. Na primer, $\frac{1}{2} = \frac{2}{4}$. Na tem nivoju bo zadostovalo, če racionalna števila definiramo kot števila, ki jih je mogoče predstaviti z ulomki. Ugotovimo, da so na množici racionalnih števil izvedljive vse štiri osnovne računske operacije z edino izjemo, da deljenje z 0 ni definirano. Ali se lahko tu ustavimo? Saj smo vendar rešili vse probleme, povezane z osnovnimi računskimi operacijami! Očitno pa nam še nekaj manjka. Zakaj bi sicer govorili o realnih številih?

Na tem mestu lahko opišemo predstavitev racionalnih števil na številski premici. Potem pa pozornost usmerimo na merjenje razdalj med točkami na številski premici. Vsako razdaljo si želimo izraziti z nekim številom. Pri tem nam bo razdalja med točkama 0 in 1 služila kot enota. Vsako točko na številski premici, ki predstavlja kako racionalno število, bomo imenovali racionalna točka. Naj točka P leži desno od izhodišča na številski premici. Če je P racionalna točka, ki predstavlja racionalno število x, potem je seveda razdalja med 0 in P enaka x. Na primer, točka $\frac{5}{3}$ je za $\frac{5}{3}$ oddaljena od izhodišča. Da bi lahko merili vse razdalje, moramo vsaki točki na številski premici desno od izhodišča prirediti število, ki bo merilo razdaljo med to točko in izhodiščem.

Če bi bila vsaka točka na številski premici racionalna, potem bi lahko razdalje merili že z racionalnimi števili. Izkaže pa se, da to ni res. Nad intervalom $[0, 1]$ narišemo enotski kvadrat in dolžino njegove diagonale s šestilom prenesemo na številsko premico. Dijaki vedo, da je dolžina te diagonale enaka $\sqrt{2}$. Torej smo dobili na številski premici točko, ki predstavlja število $\sqrt{2}$. Sedaj pa razložimo postopek dokazovanja s protislovjem in podamo dokaz, da $\sqrt{2}$ ni racionalno število. Tako smo dobili zgled števila, ki ni racionalno, a se pojavi pri merjenju razdalj. Sedaj pa opravimo še nekoliko daljši, a hkrati tudi bolj poučen razmislek, ki pove, da niso vse točke na številski premici racionalne. Najprej točke na številski premici uredimo. Za točki P in R na številski premici bomo pisali $P \leq R$, če je $P = R$ ali P leži levo od R. Zakaj ravno taka definicija urejenosti? Omejimo se na racionalne točke. Pozitivna racionalna števila merijo razdaljo od izhodišča, in ko se pomikamo v desno, postaja ta razdalja čedalje večja. Torej je relacija $\leq$ za pozitivne racionalne točke definirana na naraven način. Kako pa naj med sabo primerjamo pozitivno in negativno racionalno število ali pa dve negativni racionalni števili? Kaj je večje: 1 ali -2? -2 ali -4? Po naši definiciji urejenosti je $-2 \leq 1$ in $-4 \leq -2$, saj točka -2 leži levo od 1 in desno od -4. Da bi doumeli, da je tako definirana urejenost zares naravna, si mislimo, da racionalna števila opisujejo stanje na našem računu v banki. Smiselno je zapisati $x \leq y$, če je y za nas ugodnejše (ali enako) stanje na računu kot x. Potem pa ni več dvoma o tem, da je -2 manj kot 1 in več kot -4.

Naj bo P točka na številski premici. Tej točki bomo priredili decimalno število x. Premica je s celoštevilskimi točkami razdeljena na enotske podintervale in P bodisi leži med dvema zaporednima celoštevilskima točkama bodisi sama predstavlja eno od celih števil. V obeh primerih lahko najdemo tako celo število c_0, da velja

$$c_0 \leq P \leq c_0 + 1 \,.$$

Opazimo še, da je celo število c_0 v prvem primeru enolično določeno, v drugem pa ne. Sedaj pa interval med c_0 in $c_0 + 1$ razdelimo na deset enako dolgih podintervalov. Krajišča teh intervalov so c_0, $c_0 + \frac{1}{10}$, $c_0 + \frac{2}{10}$, $\ldots$, $c_0 + \frac{9}{10}$, $c_0 + 1$. Ponovno ugotovimo, da točka P bodisi leži med dvema zaporednima krajiščema bodisi je eno od teh krajišč. V obeh primerih obstaja taka števka c_1, da je

$$c_0 + \frac{1}{10}c_1 \leq P \leq c_0 + \frac{1}{10}c_1 + \frac{1}{10} \,.$$

Točka P je sedaj locirana na intervalu dolžine $\frac{1}{10}$. S ponovno razdelitvijo tega intervala na 10 enako dolgih podintervalov dobimo tako števko c_2, da velja

$$c_0 + \frac{1}{10}c_1 + \frac{1}{100}c_2 \leq P \leq c_0 + \frac{1}{10}c_1 + \frac{1}{100}c_2 + \frac{1}{100} \,.$$

Ta postopek ponavljamo. Po n korakih imamo

$$c_0 + \frac{1}{10}c_1 + \cdots + \frac{1}{10^n}c_n \leq P \leq c_0 + \frac{1}{10}c_1 + \cdots + \frac{1}{10^n}c_n + \frac{1}{10^n} \,,$$

kjer so $c_1, \ldots, c_n$ števke. S tem imamo točko P locirano na intervalu dolžine $\frac{1}{10^n}$.

Nadaljevanje tega procesa točki P priredi celo število c_0 in neskončno zaporedje števk c_1, c_2, c_3, $\ldots$ Velja pa tudi obratno. Vsakemu celemu številu c_0 in neskončnemu zaporedju števk c_1, c_2, c_3, $\ldots$ pripada natančno določena točka P na številski premici. Oglejmo si to bolj natančno. Naj bo dano celo število c_0 in neskončno zaporedje števk c_1, c_2, c_3, $\ldots$ Na številski premici označimo celoštevilski točki c_0 in $c_0 + 1$. Množico vseh točk T, za katere velja $c_0 \leq T \leq c_0+1$, imenujemo zaprti interval $[c_0, c_0+1]$. Označimo ga z I_0. Podobno zaprti interval $I_1 = \left[c_0 + \frac{1}{10}c_1, c_0 + \frac{1}{10}c_1 + \frac{1}{10}\right]$ definiramo kot množico vseh točk T, ki zadoščajo neenakosti

$$c_0 + \frac{1}{10}c_1 \leq T \leq c_0 + \frac{1}{10}c_1 + \frac{1}{10}\,.$$

Interval I_1 je podinterval intervala I_0. Na podoben način definiramo interval $I_2 = \left[c_0 + \frac{1}{10}c_1 + \frac{1}{100}c_2, c_0 + \frac{1}{10}c_1 + \frac{1}{100}c_2 + \frac{1}{100}\right]$ in potem še celo zaporedje vloženih intervalov I_3, I_4, ... Dolžine teh intervalov postanejo sčasoma tako majhne, kot le hočemo, in intuitivno je jasno, da obstaja natanko ena točka P na številski premici, ki leži v vsakem od teh intervalov.

Za hip se omejimo na točke na številski premici, ki ležijo desno od točke 0. Vsaki taki točki P smo priredili celo število $c_0 \geq 0$ in neskončno zaporedje števk c_1, c_2, c_3, ... Potem pa smo opravili še „obraten" proces: celemu številu $c_0 \geq 0$ in neskončnemu zaporedju števk c_1, c_2, c_3, ... smo priredili točko P na številski premici. Če uporabimo običajni decimalni zapis, smo torej poiskali zvezo med točkami desno od 0 na številski premici in neskončnimi decimalnimi števili $c_0{,}c_1c_2c_3\ldots$

Vsaki točki desno od izhodišča smo priredili neskončno decimalno število. Tu bi bila na mestu pripomba, da nekaterim točkam ustrezajo končne decimalne reprezentacije. Na primer, točki $\frac{9}{8}$ pripada končni decimalni zapis $\frac{9}{8} = 1{,}125$. Vendar imajo tudi taka števila neskončno decimalno reprezentacijo, v kateri se pač od nekje naprej ponavlja števka 0. V našem primeru je $\frac{9}{8} = 1{,}125000\ldots$

Kaj pa točke levo od izhodišča na številski premici? Tudi vsaki taki točki smo priredili celo število c_0 in neskončno zaporedje števk c_1, c_2, c_3, ... To točko smo torej predstavili z decimalnim številom

$$c_0 + 0{,}c_1c_2c_3\ldots$$

Ta zapis zlahka spremenimo v običajni decimalni zapis. Oglejmo si to na primeru. Točki $-1/3$ pripada po zgoraj opisanem postopku celo število -1 in zaporedje števk: 6, 6, 6, ... Od tod pa hitro dobimo običajni decimalni zapis tega števila:

$$-\frac{1}{3} = -1 + 0{,}666\ldots = -0{,}333\ldots$$

Intuitivno nam je sedaj jasno, da točkam na številski premici ustrezajo neskončna decimalna števila in obratno. Oglejmo si bolj podrobno decimalni zapis racionalnih števil. Začnimo z nekaj zgledi:

$$\frac{1}{5} = 0{,}2 \qquad \frac{1}{3} = 0{,}333\ldots = 0{,}\overline{3}$$

$$\frac{1}{7} = 0{,}142857142857142857\ldots = 0,\overline{142857}$$

Črta v zapisu pomeni, da se skupina števk pod njo od tam naprej ponavlja. V navedenih primerih je decimalni zapis bodisi končen ali pa je periodičen (od nekje naprej ponavljajoč se zapis). Srednješolci vedo, da imajo pozitivna racionalna števila bodisi končen bodisi periodično ponavljajoč se decimalni zapis. Vedo tudi, da če neki točki na številski premici pripada periodično ponavljajoč se decimalni zapis, potem ta točka predstavlja racionalno število.

V decimalnem številu

$$1{,}0100100010000100000100000001\ldots$$

pa se nobena skupina števk ne ponavlja periodično. Torej temu decimalnemu številu ustreza točka na številski premici, ki ni racionalna. Takim točkam pravimo iracionalne točke, ustrezna decimalna števila pa imenujemo iracionalna števila.

Skratka, če hočemo meriti razdalje, potem nam samo racionalna števila ne zadoščajo. Moramo jih razširiti z množico iracionalnih števil. S to razširitvijo dobimo množico realnih števil.

Je pa mogoče tudi zelo na hitro razložiti, da samo z racionalnimi števili v matematiki ne pridemo prav daleč. Spomnimo se cenenih kalkulatorjev, ki imajo poleg štirih osnovnih računskih operacij še tipko za računanje kvadratnega korena. Včasih smo znali tudi kvadratni koren „izračunati peš“. Je pa v primerjavi s štirimi osnovnimi računskimi operacijami kvadratni koren računska operacija, ki ni vedno izvedljiva v množici racionalnih števil (spomnimo se dokaza, da $\sqrt{2}$ ni racionalno število).

Vse do te točke nam srednješolci zlahka sledijo. Večino povedanega poznajo in razumejo. Sedaj pa lahko zastavimo vprašanje, kako definirati realna števila. „Odgovor“ je na dlani. Definiramo jih kot množico decimalnih števil.

Sedaj, ko nam je „vse jasno“, pa jih vprašamo, katero število ima decimalni zapis $0{,}9999\ldots = 0,\overline{9}$? Skupaj ugotovimo, da je to število bližje 1 kot število 0,9. Torej je oddaljeno od 1 za manj kot $\frac{1}{10}$ in tudi za manj kot $\frac{1}{100}$ in tudi za manj kot $\frac{1}{1000}$ … Potem pa že večina ugane, da gre za drugačen decimalni zapis števila 1. Lahko pa tudi računamo:

$$x = 0{,}99999\ldots$$

in zato

$$10x = 9{,}99999\ldots,$$

od koder dobimo $9x = 10x - x = 9{,}99999\ldots - 0{,}99999\ldots = 9$, od koder sledi, da je $x = 1$.

Tu je torej težava. Možno je, da ima neko število dve različni decimalni reprezentaciji. Na primer, število 1 ima neskončni decimalni zapis $1 = 1{,}000\ldots$ Pozorni bralec pa je opazil, da bi pri prej opisanem postopku točki 1 na številski premici lahko priredili tudi drugačen decimalni zapis $1 = 0{,}999\ldots$ Namreč, v prvem koraku bi si lahko izbrali namesto $c_0 = 1$ tudi $c_0 = 0$, saj bi bil tudi s to izbiro zahtevani pogoj $c_0 \leq 1 \leq c_0 + 1$ očitno izpolnjen. Ko bi potem v naslednjem koraku razdelili interval $[0, 1]$ na deset enako dolgih podintervalov, bi seveda opazili, da leži točka 1 na zadnjem od teh intervalov (točka 1 je celo skrajno desna točka zadnjega podintervala) in bi zato morali za prvo števko v pripadajočem zaporedju izbrati 9. To bi se ponovilo v vseh naslednjih korakih, in tako bi dobili decimalno reprezentacijo $1 = 0{,}999\ldots$

Nadarjeni srednješolci bi znali tudi sami razmisliti, katere so tiste točke na številski premici, ki imajo (dve) različni decimalni reprezentaciji (sedaj je jasno, zakaj smo uporabili narekovaje, ko smo prirejanju točke k danemu decimalnemu zapisu rekli „obraten“ proces k procesu prirejanja decimalnega števila k dani točki). In potem lahko skupaj ugotovimo, da je pač treba nekoliko več previdnosti. Množice realnih števil ne moremo definirati kot množico decimalnih števil, saj je treba nekatera različna decimalna števila identificirati, ker predstavljajo isto število.

Sedaj pa pride na vrsto vprašanje, ki nam povzroči resnične težave. S števili hočemo računati. Končni decimalni števili 0,985 in 32,451 znamo sešteti:

$$\begin{array}{r} 0{,}985 \\ +\ 32{,}451 \\ \hline =\ 33{,}436 \end{array}$$

Seštevanje smo izvedli „od zadaj naprej“. Kaj pa to pomeni, kadar imamo opravka z neskončnimi decimalnimi zapisi? Nekako v zadregi ugotovimo, da na to vprašanje ne znamo odgovoriti. Torej sploh ne vemo, kaj je vsota dveh realnih števil. Potolažimo se z mislijo, da računske operacije na realnih številih intuitivno razumemo. Saj tudi, ko računamo obseg kroga s polmerom 2, vemo, da $4 \cdot 3{,}14$ ni obseg tega kroga, ampak le približek. In da dobimo boljši približek, če namesto 3,14 vzamemo boljši približek števila π.

Pa vendar ne moremo biti zadovoljni. Pojem realnega števila bi radi vpeljali tako, da bi lahko nedvoumno povedali, kaj je vsota (produkt) dveh realnih števil.

Preden se lotimo matematično korektne vpeljave realnih števil, se moramo vprašati, kaj sploh hočemo. Na množici realnih števil hočemo imeti definirani računski operaciji seštevanje in množenje, ki zadoščata običajnim zakonitostim (komutativnost, asociativnost, distributivnost, obstoj ničelnega elementa in enote za množenje, obstoj nasprotnih in inverznih elementov). Skratka, hočemo, da množica realnih števil zadosti vsem aksiomom iz definicije polja. Hočemo pa še več; za množico realnih števil bomo zahtevali, da je urejeno polje. To pomeni, da hočemo imeti na množici realnih števil definirano še delno urejenost $\leq$ (tu seveda srednješolcem pojem delne urejenosti razložimo), ki je usklajena z algebraičnimi operacijami, kar pomeni, da za vsako trojico realnih števil a, b, c iz $a \leq b$ sledi $a + c \leq b + c$, iz $0 \leq a$ in $0 \leq b$ pa sledi $0 \leq ab$.

Vse to pa še ni dovolj. Namreč, tudi množica vseh racionalnih števil z običajnimi algebraičnimi operacijami in običajno urejenostjo je urejeno polje. Kaj je tisto, kar bo ločilo urejeno polje realnih števil od urejenega polja racionalnih števil? Odgovor se skriva v aksiomu Dedekindove polnosti.

Da bi ta aksiom razložili, potrebujemo pojem omejene množice. Najprej si bomo ogledali nekaj primerov, potem pa zapisali definicijo. Tu pa moramo biti previdni. Naš namen je matematično korektno vpeljati realna števila. Ta trenutek torej sploh še ne vemo, kaj realna števila so. Zato lahko delamo samo z racionalnimi števili in zato bomo seveda primere izbrali v množici racionalnih števil.

Oglejmo si množice racionalnih števil

$$A = \{r \in \mathbb{Q} \colon\ 1 \leq r \leq 2\}, \quad B = \{r \in \mathbb{Q} \colon\ 1 < r < 2\}$$

in

$$C = \{r \in \mathbb{Q} \colon\ r \leq 2\}.$$

Vsi takoj uganemo, da so vse tri množice navzgor omejene, navzdol omejeni pa sta samo množici A in B. Znamo tudi povedati, zakaj so vse tri množice navzgor omejene. Zato, ker je vsak element množice A (B, C) manjši ali kvečjemu enak 2.

Definicija 1. Naj bo X neprazna podmnožica delno urejene množice $(Y, \leq)$. Element $y \in Y$ je zgornja meja množice X, če velja $x \leq y$ za vsak $x \in X$. Množica X je navzgor omejena, če ima zgornjo mejo.

Seveda je 17 zgornja meja za vsako izmed množic A, B, $C \subset \mathbb{Q}$. Vsako racionalno število ≥ 2 je zgornja meja za vsako od teh množic. Ampak izmed vseh teh zgornjih mej se nam meja 2, ki je najmanjša med njimi, seveda zdi najboljša.

Definicija 2. Naj bo X neprazna navzgor omejena podmnožica delno urejene množice $(Y, \leq)$. Element $y_0 \in Y$ je natančna zgornja meja množice X, če je y_0 zgornja meja množice X in če za vsako zgornjo mejo y množice X velja: $y_0 \leq y$.

Natančni zgornji meji množice X pravimo tudi najmanjša zgornja meja množice X ali supremum množice X. Označimo jo s simbolom $\sup X$.

Torej imamo v zgornjih treh primerih

$$\sup A = \sup B = \sup C = 2\,.$$

Sedaj pa si poglejmo množico

$$D = \{r \in \mathbb{Q}\colon\ r \leq 0 \ \text{ ali } \ r^2 \leq 2\}\,.$$

Intuitivno nam je takoj jasno, kaj so elementi te množice. „To so ravno vsa racionalna števila, ki so manjša od $\sqrt{2}$“. Zakaj narekovaji? Vemo že, da $\sqrt{2}$ ni racionalno število. O $\sqrt{2}$ bomo lahko govorili šele potem, ko bomo realna števila skonstruirali. Seveda pa nam nihče ne brani, da bi že sedaj, ko imamo na voljo samo racionalna števila, uporabljali intuitivno razumevanje. Tako bomo od sedaj naprej razpravljali ves čas na dveh nivojih. Na intuitivnem nivoju imamo že kar dobro predstavo o tem, kaj realna števila so, na formalnem nivoju pa lahko delamo samo z racionalnimi števili.

Množica D je navzgor omejena. Racionalno število 2 je zgornja meja te množice. Pa tudi 1,5 je zgornja meja te množice. In prav tako 1,42 … Kaj pa je natančna zgornja meja te množice?

Intuitiven odgovor je jasen: to je $\sqrt{2}$. Ampak mi imamo za zdaj opravka samo z racionalnimi števili. Točko $\sqrt{2}$ smo na številski premici že skonstruirali in tudi ugotovili, da tej točki ne ustreza nobeno racionalno število. Vsa racionalna števila, ki leže desno od te točke, so zgornje meje množice D. Vendar pa med temi zgornjimi mejami ni najmanjše.

Zakaj ne? Intuitivno je odgovor jasen. Zato, ker nam manjka $\sqrt{2}$. Ko pa bomo skonstruirali realna števila, se nam kaj takega ne bo več zgodilo.

Rekli bomo, da urejeno polje zadošča Dedekindovemu aksiomu polnosti, če ima vsaka neprazna navzgor omejena podmnožica natančno zgornjo mejo. Urejeno polje, ki zadošča temu aksiomu, se imenuje Dedekindovo urejeno polje.

Kaj bo torej sedaj naša naloga? Imamo množico racionalnih števil. Skonstruirali bomo množico realnih števil, in sicer tako, da bomo po eni strani zadovoljili našo intuitivno predstavo (vsakemu realnemu številu bo ustrezala natanko ena točka na številski premici in vsaki točki na številski premici bo ustrezalo natanko eno realno število, računske operacije in urejenost bodo definirane tako, da se bodo definicije skladale z našo intuicijo), po drugi strani pa bo skonstruirana množica skupaj z na njej definiranima računskima operacijama seštevanja in množenja in z na njej definirano relacijo delne urejenosti ustrezala vsem aksiomom iz definicije Dedekindovega urejenega polja.

Kako začeti? Idejo dobimo, če se spomnimo množice D. V tem trenutku smo na intuitivni strani. Množica D je podmnožica racionalnih števil. Intuitivno pa je jasno, da nam ta množica določa število $\sqrt{2}$. In tu je ideja! Če mislimo o realnih številih kot o točkah na številski premici in če se zgledujemo po množici D, potem je vsaka točka (= vsako realno število) določena z množico vseh racionalnih točk (= racionalnih števil), ki leže levo od te točke (= levo od tega racionalnega števila). Torej bomo predstavili vsako realno število x z množico vseh racionalnih števil r, za katera velja $r < x$. Tako množico bomo imenovali Dedekindov rez.

Ponovimo: naša naloga je skonstruirati množico realnih števil. Za zdaj imamo na voljo samo racionalna števila. Torej moramo Dedekindove reze definirati samo z racionalnimi števili. Intuitivno je Dedekindov rez, ki predstavlja število x, presek poltraka $(-\infty, x)$ z množico racionalnih števil. Kako opisati take množice? Kaj je njihova karakteristična lastnost? Po nekaj premišljevanja se nam posveti, da je Dedekindov rez α taka podmnožica racionalnih števil, da hkrati z vsakim svojim elementom r vsebuje tudi vsa racionalna števila q, ki so manjša od r. Pravkar zapisano lastnost pa imata seveda tudi prazna množica in množica $\mathbb{Q}$. Torej imamo tri lastnosti; poleg zgoraj opisane sta tu še nepraznost in različnost od $\mathbb{Q}$. Tu bi morda pomislili, da bomo Dedekindove reze definirali kar kot podmnožice racionalnih števil, ki imajo te tri lastnosti. Pa bo le treba še nekaj previdnosti! Ko intuitivno premišljujemo, kaj bi bili vsi Dedekindovi rezi ob taki definiciji, se nam kmalu posveti, da bi (intuitivno!) to bili bodisi preseki poltrakov $(-\infty, x)$ z množico racionalnih števil bodisi preseki poltrakov $(-\infty, x]$ z množico racionalnih števil.[1] Da bi izločili drugo možnost, bomo od Dedekindovega reza zahtevali še, da ne vsebuje največjega elementa (seveda je največji element podmnožice $A \subset \mathbb{Q}$ definiran kot tako racionalno število $r \in A$, da je $p \leq r$ za vsak $p \in A$). Izkaže se, da smo sedaj že na cilju! In lahko začnemo s formalnimi definicijami.

Definicija 3. Dedekindov rez je podmnožica $\alpha \subset \mathbb{Q}$, ki ima naslednje lastnosti:

1. $\alpha \neq \emptyset$,
2. $\alpha \neq \mathbb{Q}$,
3. α ne vsebuje največjega elementa,
4. za vsak par r, $q \in \mathbb{Q}$ iz $r \in \alpha$ in $q < r$ sledi $q \in \alpha$.

Vsak tak Dedekindov rez nam bo predstavljal neko realno število. Torej bomo definirali:

Definicija 4. Realno število je Dedekindov rez. Množico vseh realnih števil označimo z $\mathbb{R}$.

[1] Srednješolci z dobrim intuitivnim razumevanjem bodo hitro uganili, da se tidve množici ne razlikujeta, kadar je x iracionalen.

Množico realnih števil imamo. Kako jo bomo uredili? Intuitivni razmislek je preprost. Za realni števili x, y velja $x \leq y$ natanko tedaj, ko je $(-\infty, x) \subset (-\infty, y)$. Ponovno smo razmišljali s poltraki in intervali, ki so za zdaj še nedefinirani pojmi. Ampak mi intuitivno že vemo, za kaj gre! Formalno pa zgodbo peljemo naprej takole:

Definicija 5. Za poljubni dve realni števili (za poljubna dva Dedekindova reza) α, β definiramo

$$\alpha \leq \beta \iff \alpha \subset \beta\,.$$

Sedaj imamo več možnosti. Ker sem imel za predavanje na voljo dve šolski uri, sem hotel pojasniti zgolj glavne ideje. Če bi vodil krožek, pa bi si za to tematiko vzel več časa in bi s srednješolci zares dokazal, da smo na ta način definirali relacijo delne urejenosti. Pri tem bi jih poizkusil zgolj usmerjati in jih na ta način učiti dokazovanja matematičnih trditev. Kar je seveda zelo težko! Tu pač ne smemo pozabiti svojih nerodnih začetkov, ko smo imeli težave celo s preverjanjem povsem preprostih trditev.

Preprosto je videti, da je izpolnjen aksiom Dedekindove polnosti. Naj bo $A \subset \mathbb{R}$ neprazna navzgor omejena množica. Pokazati moramo obstoj najmanjše zgornje meje. Tu si zopet pomagamo z intuicijo. Na številski premici si ponazorimo množice

$$A = \{-1, -1/2, -1/3, -1/4, \ldots\}\,,$$

$B = (-1, 0)$ in $C = [-1, 0]$. V vseh treh primerih uganemo, da je najmanjša zgornja meja enaka 0, in se pri tem spomnimo, da je število 0 predstavljeno kot množica vseh negativnih racionalnih števil. Sedaj pa si elemente množic A, B in C predstavljamo kot Dedekindove reze (poltrake, presekane s $\mathbb{Q}$). Bistri srednješolci bodo kmalu ugotovili, da je v vseh treh primerih Dedekindov rez, ki predstavlja 0, unija vseh Dedekindovih rezov, ki predstavljajo vsa števila iz A, oziroma B, oziroma C.

Trditev 1. *Naj bo $A \subset \mathbb{R}$ neprazna navzgor omejena množica. Potem je*

$$\bigcup_{\alpha \in A} \alpha$$

natančna zgornja meja množice A.

Tu ne bom navajal dokaza, ki ga zainteresirani bralec zlahka najde v literaturi ali na spletu. Pravzaprav verjamem, da je večina bralcev, ki so se prebili do te točke, sposobna sama dokazati gornjo trditev in jim bo literatura služila zgolj za preverjanje, da niso česa pomembnega izpustili.

Da bi končali s konstrukcijo realnih števil, moramo vpeljati obe računski operaciji, preveriti zahtevane lastnosti in usklajenost z relacijo urejenosti. Kako vpeljemo seštevanje? Intuicija pravi, da je realno število $\sqrt{2}$ predstavljeno z vsemi racionalnimi števili, ki so manjša od $\sqrt{2}$, da je realno število $\sqrt{3}$ predstavljeno z vsemi racionalnimi števili, ki so manjša od $\sqrt{3}$, in da je realno število $\sqrt{2}+\sqrt{3}$ predstavljeno z vsemi racionalnimi števili, ki so manjša od $\sqrt{2}+\sqrt{3}$. Če seštejemo racionalni števili $p<\sqrt{2}$ in $q<\sqrt{3}$, potem naj bi bilo $p+q<\sqrt{2}+\sqrt{3}$. Intuicija še pove, da lahko vsako racionalno število $r<\sqrt{2}+\sqrt{3}$ razcepimo v obliki $r=p+q$, kjer $p<\sqrt{2}$ in $q<\sqrt{3}$. Od intuitivnega gremo nazaj k formalnemu pristopu in zapišemo:

Definicija 6. Za poljubna Dedekindova reza α, β definiramo

$$\alpha+\beta=\{p+q\colon\, p\in\alpha \ \text{ in } \ q\in\beta\}\,.$$

No, sedaj ni težko preveriti, da je za vsak par α, $\beta\in\mathbb{R}$ tudi $\alpha+\beta$ Dedekindov rez in da tako definirano seštevanje izpolnjuje vse zahtevane lastnosti.

Pri definiciji množenja je treba biti nekoliko bolj previden. Tu bomo izpustili podrobnosti. Naš namen je bil opisati, kako je mogoče nadarjenim srednješolcem predstaviti zelo zahtevno matematično tematiko. Če bi kak mentor v srednji šoli izbral to tematiko za serijo predavanj v okviru krožkov, bo po do sedaj povedanem zlahka našel ustrezno literaturo ali zapise na spletu ali pa celo stare zapiske iz Analize 1, kjer se bo lahko poučil o manjkajočih detajlih in si ob pomoči tega zapisa izbral svoj koncept predstavitve te zahtevne (a hkrati intuitivno dokaj jasne) tematike.

Kaj smo počeli? Začeli smo z urejenim poljem $\mathbb{Q}$. Potem smo samo z uporabo racionalnih števil skonstruirali množico realnih števil in na njej definirali seštevanje, množenje in urejenost. Pri tem smo izpolnili vse aksiomatske zahteve in tudi intuitivno se tako definirana realna števila natanko ujemajo z našo predstavo. Seveda tu povemo še, kako je vpeljavo realnih števil običajno videti v literaturi. Tu smo (opisoval sem predavanje) imeli ves čas opravka z vzporednima zgodbama: intuitivno in strogo formalno. Na krožku, kjer lahko predavanje raztegnemo na več ur, bi se spodobilo še enkrat ponoviti zgolj povsem formalno vpeljavo realnih števil brez kakršnihkoli intuitivnih vložkov. In tu lahko spregovorimo besedo ali dve o razmerju med intuitivnim dojemanjem in matematično strogostjo. Intuitivno dojemanje je bistveno, matematična strogost pa je nujno potrebna. Razvoj vseh znanosti se včasih bere skoraj kot opis zaporedja znanstvenih zmot. Do neke mere je to res celo za matematiko. A je pri izogibanju napakam matematika uspešnejša od drugih znanosti. Tudi zato, ker sledimo pravilom matematične strogosti, kot smo storili tudi pri vpeljavi realnih števil. Zavedati se moramo, da se zgolj pri sklicevanju na intuicijo, zlasti v bolj zapletenih situacijah, napakam težko izognemo.

Sedaj pa nam je jasno, da smo se na začetku le prenaglili, ko smo „ugotovili“, da vemo vse o realnih številih. Ampak težave smo odpravili in sedaj le kaže, da je ta zgodba končana.

Pa seveda to še zdaleč ni res! Realna števila smo skonstruirali in tako dobili matematični objekt, ki zadošča vsem aksiomom za Dedekindovo urejeno polje. Vendar se tu takoj pojavi še vprašanje edinosti. Učeno rečemo, da je s temi aksiomi Dedekindovo urejeno polje do izomorfizma enolično določeno. Srednješolcem, ki so sledili razlagi do te točke, je mogoče intuitivno razložiti pomen tega stavka. In potem lahko nadaljujemo in povemo, da tudi s tem zgodbe še ni konec. Ali je mogoče realna števila skonstruirati še na kak drug naraven način? Seveda je to mogoče, na primer z napolnitvijo polja racionalnih števil. Ali zares potrebujemo vse aksiome, ki smo jih našteli? Drugače rečeno, ali lahko katerega od aksiomov izpustimo, ne da bi s tem ogrozili enolično določenost (do izomorfizma) polja $\mathbb{R}$? Ali obstaja še kakšna druga naravna aksiomatizacija realnih števil? Obstajajo, na primer Tarskijeva aksiomatizacija. Podrobnejšo razlago bo zainteresirani bralec zlahka našel na spletu (angleška izraza sta „completion of the rational numbers“ in „Tarski's axiomatization of the reals“). In še opomba čisto na koncu: realna števila smo skonstruirali z racionalnimi števili. Tu so stvari dokaj jasne. Vemo, kateri ulomki predstavljajo isto racionalno število, vemo, kako racionalna števila (ulomke) seštevamo in množimo in kako definiramo delno urejenost $\leq$. No, če hočemo imeti matematično povsem korektno teorijo številskih množic $\mathbb{N}$, $\mathbb{Z}$ in $\mathbb{Q}$, pa je tudi tu treba še marsikaj povedati.

Skupaj s srednješolci se lahko na koncu vprašamo, čemu vse to služi. Matematiko uporabljamo vsepovsod, pogosto v navezi z računalniki. Takrat seveda računamo z racionalnimi števili (končnimi decimalnimi števili). Tako je na prvi pogled videti, da bi lahko shajali brez realnih števil. Zdi se, da bi zadostovala racionalna števila. Še posebej pa se vprašanja, zastavljena v prejšnjem odstavku, marsikomu zdijo odvečna.

Del odgovora je, da je matematika zelo uporabna veda. Pri njeni uporabi se moramo zavedati, kaj je zagotovo res in kaj ne. Poznati moramo torej teorijo (izreke). Da pa bi izgradili teorijo, nujno potrebujemo matematično korektno vpeljano množico realnih števil, saj se sicer zatakne že pri najosnovnejših pojmih.

Seveda pa je v igri še vse kaj drugega. Vprašanja, ki smo se jih tu lotili, preprosto zahtevajo odgovor. Večine ljudi ta vprašanja ne zanimajo. A tako, kot se bodo vedno našli ljudje, ki bodo morali pisati pesmi, morali slikati, morali ustvarjati glasbo . . . se bodo vedno našli tudi taki, ki bodo preprosto morali vedeti ali vsaj iskati odgovore na taka in podobna vprašanja. Ukvarjanje s filozofijo, umetnostjo, matematiko . . . je pač nepogrešljivi del človeške narave.

ŠTIRINAJSTO MEDNARODNO TEKMOVANJE ŠTUDENTOV MATEMATIKE

Med 3. in 9. avgustom 2007 je v Blagoevgradu v Bolgariji potekalo že štirinajsto mednarodno tekmovanje študentov matematike. Prva tekmovanja so bila organizirana v okviru projekta Tempus in so bila namenjena primerjavi kvalitete študija na evropskih univerzah. Zadnjih nekaj let pa prihaja tekmovat tudi čedalje več študentov iz Amerike in Azije in tekmovanje počasi postaja študentski ekvivalent srednješolskih matematičnih olimpijad. Rastoči ugled tekmovanja se kaže tudi v dejstvu, da so najboljšim tekmovalcem nekatere tuje univerze ponudile polne štipendije za dodiplomski in podiplomski študij pri njih. Za bodoče kadre so se zelo opazno potegovale tudi finančne inštitucije.

Slika 1. Slovenska ekipa: Nik Stopar, Sara Kališnik, Uroš Kuzman in Kris Stopar. Čepi vodja ekipe Marjan Jerman.

Iz prejšnjih let vsega hudega navajeni tekmovalci in vodje ekip smo bili pozitivno presenečeni nad zahodnimi standardi tekmovanja in bivanja, ki sta bila organizirana v okviru Ameriške univerze v Bolgariji. Blagoevgrad je eno najlepših in najbolj urejenih bolgarskih mest. Po tekmovanju so si študentje ogledali mesto Bansko, ki je pravi naravni biser in bolgarsko smučarsko središče; bolj radovedni pa so obiskali še bližnji znameniti rilski samostan.

Ljubljansko univerzo so zastopali Sara Kališnik iz drugega, Kris Stopar in Nik Stopar iz tretjega ter Uroš Kuzman iz četrtega letnika.

Naloge so bile letos težje kot po navadi. Le 50 (od 249) študentov je doseglo več kot polovico možnih točk. Nik Stopar je prejel drugo nagrado, Uroš Kuzman pa je dobil pohvalo. Tradicionalno je vsaka ekipa, ki predstavlja univerzo, sestavljena iz štirih študentov, nekatere univerze pa na tekmovanje pošljejo tudi večje ekipe. Letos je bil prvič objavljen tudi vrstni red ekip, kjer so za merilo vzeli vsoto točk najbolje uvrščenih štirih študentov iz ekipe. Prva tri mesta so zavzeli Madžari, Rusi in Iranci. Slovenska ekipa je zasedla 33. mesto med šestdesetimi univerzami.

Tako kot vsa leta je tekmovanje potekalo v prijetnem in prijateljskem ozračju. Nenavadno deževno in hladno vreme je preprečilo že skoraj tradicionalno tekmovanje ekip v nogometu.

Dva dneva so študentje vsak dan po pet ur reševali po šest nalog. Teža nalog približno narašča z zaporedno številko. Praviloma je prva naloga namenjena ogrevanju, zadnja pa je zelo težka.

Radoveden bralec lahko vseh dvanajst nalog najde na študentskih straneh Fakultete za matematiko in fiziko. Na tem mestu navedimo le štiri. Bralce vabim, da jih poskušajo rešiti sami. Za boljši uvid v težo nalog objavljamo tudi njihove rešitve.

I.3. Za polinom p v spremenljivkah $x_1, \ldots, x_k$ pravimo, da je *dober*, če obstajajo realne matrike $A_1, \ldots, A_k$ velikosti 2×2, za katere velja

$$p(x_1, \ldots, x_k) = \det\left(\sum_{i=1}^{k} x_i A_i\right).$$

Poišči vsa naravna števila k, za katera so vsi homogeni polinomi v k spremenljivkah stopnje 2 dobri.

I.6. Za polinom p s celimi koeficienti velja $|p(z)| \leq 2$ za vsa kompleksna števila z, ki ležijo na enotskem krogu. Kolikšno je največje število neničelnih koeficientov polinoma p?

II.1. Naj bo $f\colon \mathbb{R} \to \mathbb{R}$ zvezna funkcija, za katero velja: za poljubno realno število $c > 0$ lahko graf funkcije cf dobimo le s togim premikom (to je, z zaporedjem rotacij in translacij) grafa funkcije f. Ali je graf funkcije f nujno premica?

II.3. Naj bo C neprazna, zaprta in omejena podmnožica realne osi, $f\colon C \to C$ pa nepadajoča zvezna funkcija (to je, $f(x) \leq f(y)$ za vse $x \leq y$). Pokaži, da obstaja točka $p \in C$, za katero je $f(p) = p$.

Slika 2. Takole je bilo videti popravljanje naloge II.3, ki je trajalo do jutranjih ur. Vsako rešitev, ki mora biti napisana v angleščini (vsak tekmovalec je skrit pod šifro), neodvisno pregledata vsaj dve vodji ekip in jo ocenita po vnaprej določenem kriteriju. Ko je popravljanje končano, morajo popravljavci uskladiti ocene. V naslednjih dneh sledi še prizadevanje vodij ekip za višje ocene, s katerimi se morajo strinjati vsaj trije popravljavci. Na sliki je mednarodno pisana ekipa, spomnim se Brazilca, Madžara, Poljaka, Slovaka, Nizozemca in Ukrajinke.

Ne pokukajte prehitro v rešitve:

I.3. V primeru ene spremenljivke so edini homogeni polinomi stopnje dva oblike $p(x) = ax^2$ za primerno realno število a. Tedaj lahko vzamemo za $A_1 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & a \end{smallmatrix}\right]$.

Vsi homogeni polinomi stopnje dva v dveh spremenljivkah so oblike $p(x, y) = ax^2 + by^2 + cxy$. Ker je

$$p(x, y) = \begin{vmatrix} x & by \\ -y & ax + cy \end{vmatrix},$$

enakosti v nalogi zadoščata matriki $A_1 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & a \end{smallmatrix}\right]$ in $A_2 = \left[\begin{smallmatrix} 0 & b \\ -1 & c \end{smallmatrix}\right]$.

Sedaj privzemimo, da je $k \geq 3$. Pokažimo, da polinom $p(x_1, \ldots, x_k) = \sum\limits_{i=1}^{k} x_i^2$ ni dober. V nasprotnem primeru bi za vse izbire x_1, ..., x_k veljalo

$p(x_1, \ldots, x_k) = \det\left(\sum_{i=1}^{k} x_i A_i\right)$. Prvi stolpci treh 2×2 matrik so zagotovo linearno odvisni, zato lahko izberemo takšna števila $y_1, \ldots, y_k$, da je prvi stolpec netrivialne linearne kombinacije $y_1 A_1 + \ldots + y_k A_k$ enak $\left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$. V tem primeru je $\det(y_1 A_1 + \ldots + y_k A_k) = 0$, hkrati pa je $p(y_1, \ldots, y_k) = y_1^2 + \ldots + y_k^2 \neq 0$.

I.6. Pokazali bomo, da ima lahko takšen polinom največ dva neničelna koeficienta. Pogoj, recimo, izpolnjujejo polinomi 0, 1 in $1 + z$.

Privzemimo, da tudi polinom $p(z) = a_n z^n + \ldots + a_0$, z več kot dvema neničelnima koeficientoma, izpolnjuje zahtevane pogoje. Ker lahko brez škode za splošnost polinom p delimo s primerno potenco z in morda zamenjamo p z $-p$ (pri tem se ne spremeni število neničelnih koeficientov in absolutna vrednost polinoma na enotski krožnici), lahko privzamemo, da je neničeln že prosti člen $a_0 > 0$.

Naj bo $q(z) = a_1 z + \ldots + a_{n-1} z^{n-1}$. Pokazali bomo, da je $q = 0$.

Na enotski krožnici izberimo števila $w_0, \ldots, w_{n-1}$, za katera je $a_n w_n^k = |a_n|$, to je

$$w_k = \begin{cases} e^{2k\pi i/n}, & \text{če je } a_n > 0\,; \\ e^{(2k+1)\pi i/n}, & \text{če je } a_n < 0\,. \end{cases}$$

Tedaj je

$$\sum_{k=0}^{n-1} q(w_k) = \sum_{k=0}^{n-1} q\left(w_0 e^{2k\pi i/n}\right) = \sum_{j=1}^{n-1} a_j w_0^j \sum_{k=0}^{n-1} \left(e^{2j\pi i/n}\right)^k = 0\,.$$

Zadnja vsota je enaka 0, ker si lahko števila $\left(e^{2j\pi i/n}\right)^k$ predstavljamo kot oglišča pravilnega n-kotnika, njihovo vsoto pa kot vsoto krajevnih vektorjev oglišč. Prav tako bi jo lahko izračunali tudi kot vsoto členov končnega geometrijskega zaporedja.

Izračun povprečne vrednosti polinoma p v točkah w_k nam da

$$\tfrac{1}{n} \sum_{k=0}^{n-1} p(w_k) = \tfrac{1}{n} \sum_{k=0}^{n-1} (a_0 + q(w_k) + a_n w_k^n) = a_0 + |a_n|\,,$$

zato je

$$2 \geq \tfrac{1}{n} \sum_{k=0}^{n-1} |p(w_k)| \geq \left| \tfrac{1}{n} \sum_{k=0}^{n-1} p(w_k) \right| = a_0 + |a_n| \geq 2\,.$$

Odtod sledi, da je $a_0 = |a_n| = 1$ in $|p(w_k)| = |2 + q(w_k)| = 2$ za vse k. Vse vrednosti $q(w_k)$ torej ležijo na krožnici $|z - (-2)| = 2$, njihova vsota pa je enaka 0. To je možno le, če je $q(w_k) = 0$ za vse k. To pa pomeni, da ima polinom q, ki je stopnje največ $n-1$, vsaj n različnih ničel. Zato je $q = 0$ in ima polinom $p(z) = a_0 + a_n z^n$ največ dva neničelna koeficienta.

Kot zanimivost povejmo še, da je vodja madžarske ekipe Geza Kos opazil, da se lahko rešitvi naloge približamo s pomočjo Parsevalove enakosti. Drugo normo polinoma lahko izračunamo na dva načina, s pomočjo koeficientov in z integralom kvadrata absolutne vrednosti po enotski krožnici:

$$|a_0|^2 + \cdots + |a_n|^2 = \tfrac{1}{2\pi}\int_0^{2\pi} |p(e^{it})|^2 dt \leq \tfrac{1}{2\pi}\int_0^{2\pi} 4\, dt = 4\,.$$

Odtod sledi, da so lahko največ štirje koeficienti polinoma neničelni in enaki ± 1. Kljub precejšnjim naporom pa nam naloge, tudi ob upoštevanju te omejitve, na tak način ni uspelo rešiti do konca.

II.1. Ker velja $ce^x = e^{x+\log c}$, temu pogoju zadošča tudi eksponentna funkcija $f(x) = e^x$.

II.3. Pa recimo, da je $f(x) \neq x$ za vse $x \in C$. Naj bo $[a,b]$ najmanjši zaprt interval, ki vsebuje C. Zaradi zaprtosti množice C mora veljati $a, b \in C$. Ker slika funkcija f nazaj v C in nima negibne točke, mora veljati $f(a) > a$ in $f(b) < b$. Sedaj uporabimo standardni trik iz Analize 1 in definirajmo

$$p = \sup\{x \in C;\ f(x) > x\}\,.$$

Ker je množica C zaprta, funkcija f pa je zvezna, je $f(p) \geq p$, zato je $f(p) > p$. Za vse $x \in C$, ki so večji od p, pa velja $f(x) < x$. Tako pridemo do protislovja $f(f(p)) < f(p)$ z dejstvom, da je f nepadajoča funkcija.

Marjan Jerman

STROKOVNA EKSKURZIJA DMFA V IDRIJO
Ogled tehniške dediščine (10. ali 17. maja 2008)

Ogledali si bomo muzejske zbirke v gradu Gewerkenegg, kanomeljske klavže, Antonijev rov, idrijsko kamšt, stroje ob jašku Frančiške, rudarsko hišo z etnološko zbirko in še kakšne znamenitosti Idrije.

DMFA želi obnoviti strokovne ekskurzije. Naslednja bo predvidoma septembra v inštitute v okolici Trsta. Če se zanimate za ekskurzijo v Idrijo ali za naslednje ekskurzije, se oglasite na naslov Mitja.Rosina@ijs.si, da vam bomo pošiljali nadaljnja obvestila. Poglejte tudi na spletno stran http://www.dmfa.si/.

Mitja Rosina

OBZORNIK ZA MATEMATIKO IN FIZIKO
LJUBLJANA, MAREC 2008
Letnik 55, številka 2
ISSN 0473-7466, UDK 51 + 52 + 53

VSEBINA

CONTENTS

Na naslovnici vidimo instalacijo mikrovalovnih resonatorjev, ki so bistveni sestavni del vsakega modernega pospeševalnika elektronov z visokim izkoristkom. Brez napredka superprevodne tehnologije teh resonatorjev sodobnih meritev lastnosti protonov ne bi mogli opravljati (glej članek na strani 54). Fotografija je iz arhiva Thomas Jefferson National Accelerator Facility (`http://www.jlab.org/`).